# REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

| 1. REPORT DATE (DD-MM-YYYY) | 2. REPORT TYPE | 3. DATES COVERED (From - To) |
|---|---|---|
| 27-05-2015 | Final Report | 1-Mar-2014 - 29-Feb-2016 |

| 4. TITLE AND SUBTITLE | 5a. CONTRACT NUMBER |
|---|---|
| Final Report: An Online Biosensor for the Protection of Water Supplies | W911NF-14-2-0032 |
| | 5b. GRANT NUMBER |
| | 5c. PROGRAM ELEMENT NUMBER |

| 6. AUTHORS | 5d. PROJECT NUMBER |
|---|---|
| Jeff Hasty, Scott Cookson, Michael Ferry, Ramon Huerta, Leo Baumgart, Garrett Graham, Ryan Johnson | 5e. TASK NUMBER |
| | 5f. WORK UNIT NUMBER |

| 7. PERFORMING ORGANIZATION NAMES AND ADDRESSES | 8. PERFORMING ORGANIZATION REPORT NUMBER |
|---|---|
| University of California - San Diego<br>Office of Contract & Grant Adm<br>9500 Gilman Drive, 0934<br>La Jolla, CA          92093 -0934 | |

| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) | 10. SPONSOR/MONITOR'S ACRONYM(S)<br>ARO |
|---|---|
| U.S. Army Research Office<br>P.O. Box 12211<br>Research Triangle Park, NC 27709-2211 | 11. SPONSOR/MONITOR'S REPORT NUMBER(S)<br>65405-CH-DRP.1 |

## 12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for Public Release; Distribution Unlimited

## 13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

## 14. ABSTRACT

This report marks the conclusion of a one-year project that combined Next Generation Sequencing, genetic engineering, and microfluidic technology to precisely engineer a highly sensitive and specific biosensor platform that can continuously monitor water supplies for the presence of chemical toxins. We applied synthetic biology to develop novel microbial sensor strains that have sensitive and specific responses to critical water toxins, and we used state-of-the-art microfluidic techniques and optical technology along with computational biology to detect and interpret the signals from these organisms. Our final deliverable is a self-contained water sensor prototype.

## 15. SUBJECT TERMS

synthetic biology, microfluidics, whole-cell biosensor, water contamination

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 15. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | UU | | Jeff Hasty |
| UU | UU | UU | | | 19b. TELEPHONE NUMBER<br>858-822-3442 |

## Report Title

Final Report: An Online Biosensor for the Protection of Water Supplies

## ABSTRACT

This report marks the conclusion of a one-year project that combined Next Generation Sequencing, genetic engineering, and microfluidic technology to precisely engineer a highly sensitive and specific biosensor platform that can continuously monitor water supplies for the presence of chemical toxins. We applied synthetic biology to develop novel microbial sensor strains that have sensitive and specific responses to critical water toxins, and we used state-of-the-art microfluidic techniques and optical technology along with computational biology to detect and interpret the signals from these organisms. Our final deliverable is a self-contained water sensor prototype.

## Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

### (a) Papers published in peer-reviewed journals (N/A for none)

Received          Paper

**TOTAL:**

**Number of Papers published in peer-reviewed journals:**

### (b) Papers published in non-peer-reviewed journals (N/A for none)

Received          Paper

**TOTAL:**

**Number of Papers published in non peer-reviewed journals:**

### (c) Presentations

## Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

<u>Received</u>          <u>Paper</u>

**TOTAL:**

**Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):**

## Peer-Reviewed Conference Proceeding publications (other than abstracts):

<u>Received</u>          <u>Paper</u>

**TOTAL:**

**Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):**

## (d) Manuscripts

<u>Received</u>          <u>Paper</u>

**TOTAL:**

**Number of Manuscripts:**

## Books

Received        Book

    **TOTAL:**

Received        Book Chapter

    **TOTAL:**

## Patents Submitted

An Online Biosensor for the Protection of Water Supplies

## Patents Awarded

## Awards

## Graduate Students

| NAME | PERCENT_SUPPORTED | Discipline |
|---|---|---|
| Leo A Baumgart | 0.45 | |
| **FTE Equivalent:** | **0.45** | |
| **Total Number:** | **1** | |

## Names of Post Doctorates

| NAME | PERCENT_SUPPORTED |
| --- | --- |
| Scott W Cookson | 1.00 |
| Michael S Ferry | 0.71 |
| Irene Rodriguez Lujan | 1.00 |
| **FTE Equivalent:** | **2.71** |
| **Total Number:** | **3** |

## Names of Faculty Supported

| NAME | PERCENT_SUPPORTED | National Academy Member |
| --- | --- | --- |
| Jeff M Hasty | 0.20 | No |
| Ramon Huerta | 0.25 | No |
| Lev S Tsimring | 0.17 | No |
| **FTE Equivalent:** | **0.62** | |
| **Total Number:** | **3** | |

## Names of Under Graduate students supported

| NAME | PERCENT_SUPPORTED | Discipline |
| --- | --- | --- |
| Ellixis A Julio | 0.75 | Biosciences |
| **FTE Equivalent:** | **0.75** | |
| **Total Number:** | **1** | |

## Student Metrics
This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: ...... 0.00

The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields:...... 0.00

Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale):...... 0.00

Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering:...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense ...... 0.00

The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields:...... 0.00

## Names of Personnel receiving masters degrees

| NAME |
| --- |
| **Total Number:** |

## Names of personnel receiving PHDs

| NAME |
| --- |
| **Total Number:** |

## Names of other research staff

| NAME | PERCENT_SUPPORTED |
|------|-------------------|
| Youngmin Jeon | 0.80 |
| Ryan A Johnson | 0.80 |
| **FTE Equivalent:** | **1.60** |
| **Total Number:** | **2** |

## Sub Contractors (DD882)

1 a. Ziva Corporation

1 b. 6160 Lusk Blvd.

C-206

San Diego          CA          921212740

**Sub Contractor Numbers (c):**

**Patent Clause Number (d-1):**

**Patent Date (d-2):**

**Work Description (e):** Designed a custom imaging system with appropriate size, field of view, and spectral chan

**Sub Contract Award Date (f-1):**

**Sub Contract Est Completion Date(f-2):**

1 a. Ziva Corporation

1 b. 6160 Lusk Blvd.

C-206

San Diego          CA          921212740

**Sub Contractor Numbers (c):**

**Patent Clause Number (d-1):**

**Patent Date (d-2):**

**Work Description (e):** Designed a custom imaging system with appropriate size, field of view, and spectral chan

**Sub Contract Award Date (f-1):**

**Sub Contract Est Completion Date(f-2):**

## Inventions (DD882)

# Scientific Progress

Major Developments in Year 1:

There are two major aspects to the development of our biosensor: the application of synthetic biology to develop novel microbial sensor strains that will have sensitive and specific responses to critical water toxins, and the use of state-of-the-art microfluidic techniques and optical technology along with computational biology to detect and interpret the signals from these toxin-sensing organisms.

Task A focuses mainly on the biological aspect of this project, with the goal of identifying combinations of cellular signals that can be harnessed to provide specific responses to the presence of a range of potential water toxins. In Q1, we searched the literature to identify known cellular signaling pathways responsive to our toxins of interest and selected several candidate promoters from a variety of microbial organisms. We designed plasmids with each of these promoters driving GFP and contracted these sequences to be constructed. As proof of principle, we used microfluidics to test two such plasmids that we built in-house. We subjected these preliminary sensor strains to various toxin levels within a novel microfluidic chip, and we observed bright response signals. In addition to taking advantage of known toxin-sensitive pathways, we conducted a program of Next Generation Sequencing to greatly expand the number of known response promoters for each toxin. In Q2, we developed novel RNA-Seq analysis algorithms to identify specific differentially expressed genes in our large data set. We located the promoter regions of the most promising differentially expressed genes and designed sensor circuits based on them. In Q3, we completed the construction of plasmid-based microbial sensor strains for all toxins, based on promoters identified via literature searches and RNA-Seq. We used microfluidics to demonstrate the proper induction of each strain by various levels of the relevant toxin.

Task B focuses primarily on mechanical sensor development, including the microfluidic device design, optical technology, and computational tools required to translate a series of optical signals from multiple sensor strains into a meaningful toxin level determination. In Q1, we developed a novel microfluidic device to culture and sequentially expose an array of sensing strains to various levels of our toxins of interest over a period of several weeks. This "gill" chip contains tall cell traps that provide a bright fluorescent signal from a large population of cells. We also developed complimentary technologies that allowed us to mix concentrated media with a natural water source and to dispose of the microbial species safely upon exiting the device. In Q2, we partnered with the Ziva Corporation to develop a low cost, field capable optical system to image our microfluidic devices. We demonstrated that the capabilities of this optical system were comparable to a research grade microscope at low optical power. To enable viable long-term storage of biosensor chips, we developed a method to deposit and freeze-dry strains in a defined array within our microfluidic device. We demonstrated the successful revival of our strains after four weeks of room temperature storage with little loss of viability. Lastly, we investigated replacing our GFP reporter with bacterial luciferase toward the goal of increasing SNR by eliminating background autofluorescence. We demonstrated that our arsenic sensor modified in this manner shows much greater sensitivity, even compared to analytical methods approved by the EPA for detecting arsenic in drinking water. In Q3, we used strain response data to train our classifier to identify the presence of each toxin of interest in a water source of unknown composition in real time. In Q4, we strengthened classifier performance by acquiring long data sets (up to 50 days) of both on- and off-target toxin exposures within our microfluidic devices.

Task C aims to develop technologies required to make an inexpensive and robust sensor prototype that can be deployed in real bodies of water to continuously monitor for toxin contamination. In Q1, we developed a peristaltic pumping system to mix water and growth media on-chip and a filter to prevent clogs in the microfluidic device. In Q2, we developed a comprehensive software package to control fluid flow, image acquisition, and wireless data transfer. We also constructed a solar energy harvesting system for powering our device in an off-grid field environment. In Q3 and Q4, we completed the development of a functional prototype of our biosensor for use in comprehensively testing our sensor strains. We integrated the biological, fluidic, and computational components into a temperature-controlled enclosure containing all optics and electronics necessary to draw a water sample, acquire images of the biosensor strains within our microfluidic device, analyze this data using the embedded classifier algorithm, and transmit the results via encrypted Wi-Fi.

Our final deliverable is a self-contained water sensor prototype. Attached are the Year 1 results that demonstrate our ability to acquire and analyze data in real time to provide an accurate and continuous determination of water quality.

# Technology Transfer

# Hasty - UCSD ARO Final Progress Report

**Agreement No. W911NF-14-2-0032**

**DARPA Living Foundries Grant Hasty March 2015 Progress Report**

**Agreement Number:** W911NF-14-2-0032
**Name of Recipient:** UCSD
**Project Title:** An Online Biosensor for the Protection of Water Supplies
**Project PI:** Jeff Hasty
**Report Date:** 1 Mar 2015
**Reporting Phase:** Q4
**Reporting Period:** 1 Mar 2014–1 Mar 2015

## Major developments:

There are two major aspects to the development of our biosensor: the application of synthetic biology to develop novel microbial sensor strains that will have sensitive and specific responses to critical water toxins, and the use of state-of-the-art microfluidic techniques and optical technology along with computational biology to detect and interpret the signals from these toxin-sensing organisms.

Task A focuses mainly on the biological aspect of this project, with the goal of identifying combinations of cellular signals that can be harnessed to provide specific responses to the presence of a range of potential water toxins. In Q1, we searched the literature to identify known cellular signaling pathways responsive to our toxins of interest and selected several candidate promoters from a variety of microbial organisms. We designed plasmids with each of these promoters driving GFP and contracted these sequences to be constructed. As proof of principle, we used microfluidics to test two such plasmids that we built in-house. We subjected these preliminary sensor strains to various toxin levels within a novel microfluidic chip, and we observed bright response signals. In addition to taking advantage of known toxin-sensitive pathways, we conducted a program of Next Generation Sequencing to greatly expand the number of known response promoters for each toxin. In Q2, we developed novel RNA-Seq analysis algorithms to identify specific differentially expressed genes in our large data set. We located the promoter regions of the most promising differentially expressed genes and designed sensor circuits based on them. In Q3, we completed the construction of plasmid-based microbial sensor strains for all toxins, based on promoters identified via literature searches and RNA-Seq. We used microfluidics to demonstrate the proper induction of each strain by various levels of the relevant toxin.

Task B focuses primarily on mechanical sensor development, including the microfluidic device design, optical technology, and computational tools required to translate a series of optical signals from multiple sensor strains into a meaningful toxin level determination. In Q1, we developed a novel microfluidic device to culture and sequentially expose an array of sensing strains to various levels of our toxins of interest over a period of several weeks. This "gill" chip contains tall cell traps that provide a bright fluorescent

signal from a large population of cells. We also developed complimentary technologies that allowed us to mix concentrated media with a natural water source and to dispose of the microbial species safely upon exiting the device. In Q2, we partnered with the Ziva Corporation to develop a low cost, field capable optical system to image our microfluidic devices. We demonstrated that the capabilities of this optical system were comparable to a research grade microscope at low optical power. To enable viable long-term storage of biosensor chips, we developed a method to deposit and freeze-dry strains in a defined array within our microfluidic device. We demonstrated the successful revival of our strains after four weeks of room temperature storage with little loss of viability. Lastly, we investigated replacing our GFP reporter with bacterial luciferase toward the goal of increasing SNR by eliminating background autofluorescence. We demonstrated that our arsenic sensor modified in this manner shows much greater sensitivity, even compared to analytical methods approved by the EPA for detecting arsenic in drinking water. In Q3, we used strain response data to train our classifier to identify the presence of each toxin of interest in a water source of unknown composition in real time. In Q4, we strengthened classifier performance by acquiring long data sets (up to 50 days) of both on- and off-target toxin exposures within our microfluidic devices.

Task C aims to develop technologies required to make an inexpensive and robust sensor prototype that can be deployed in real bodies of water to continuously monitor for toxin contamination. In Q1, we developed a peristaltic pumping system to mix water and growth media on-chip and a filter to prevent clogs in the microfluidic device. In Q2, we developed a comprehensive software package to control fluid flow, image acquisition, and wireless data transfer. We also constructed a solar energy harvesting system for powering our device in an off-grid field environment. In Q3 and Q4, we completed the development of a functional prototype of our biosensor for use in comprehensively testing our sensor strains. We integrated the biological, fluidic, and computational components into a temperature-controlled enclosure containing all optics and electronics necessary to draw a water sample, acquire images of the biosensor strains within our microfluidic device, analyze this data using the embedded classifier algorithm, and transmit the results via encrypted Wi-Fi.

Our final deliverable is a self-contained water sensor prototype. Below, we present an overview of the Year 1 results that demonstrate our ability to acquire and analyze data in real time to provide an accurate and continuous determination of water quality.

### Problems/Issues:

None at this time.

### Accomplishment of Year 1 Goals:

This report marks the conclusion of a one year project that combined Next Generation Sequencing, genetic engineering, and microfluidic technology to precisely engineer a highly sensitive and specific biosensor platform that can continuously monitor water supplies for the presence of chemical toxins. The following reporting update section will describe each of the project's milestones along with related data and results demonstrating their achievement. In this section we present an overview of the most relevant results
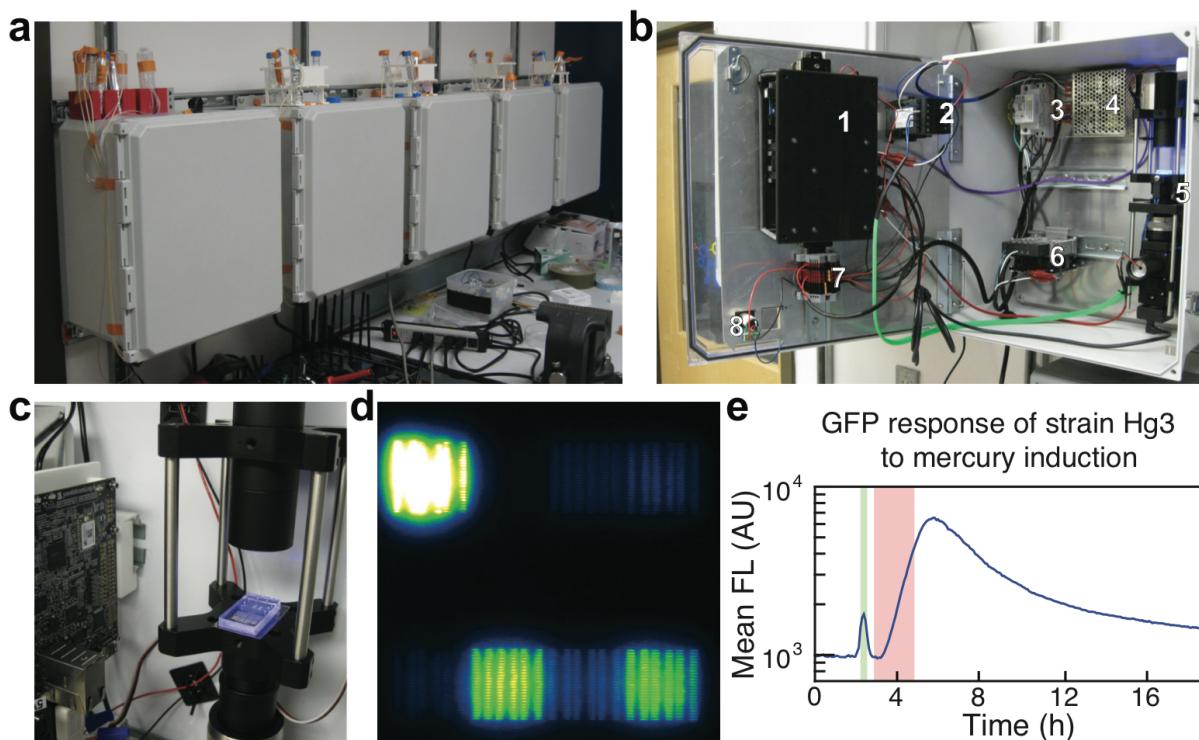
*Figure 1: An overview of our sensor prototype. a) We currently have five replicates of the sensor prototype running continuously in the lab. b) Sensor enclosure with front panel opened to expose internal components. b1: Electronics sub-enclosure, b2: temperature controller, b3: AC power distribution devices, b4: tri-output DC power supply, b5: Ziva optical assembly, b6: fan/heater, b7: DC power distribution block, b8: peristaltic pump. c) Microfluidic device being illuminated with the GFP excitation LED. d) Image taken using Ziva optics of eight strains growing inside our microfluidic device and responding to induction. e) Representative time series of data taken from one of the sensors. The green bar represents the introduction of tracer dye, and the red bar represents the introduction of the inducer, in this case mercury. Fluorescence of the mercury-specific strain rises and then falls as mercury is introduced and removed.*

that demonstrate the accomplishment of the ultimate deliverable: a finalized prototype of a biosensor for the protection of water supplies.

An overview of the sensor can be seen in Figure 1. In order to acquire the large amount of data required to successfully train the classifier, we built five replicates of the fully self-contained device prototype, in which the microfluidic devices are mounted and imaged using custom-designed optics (see Milestone 6.9 below). The majority of the data for this project was acquired using a novel microfluidic device capable of housing eight different bacterial strains that receive media from a shared source. This enables the simultaneous testing of the responses of eight different strains to each toxin concentration, providing information on specific and non-specific responses, which can be used to strengthen the classifier (Fig. 1d). We have recently expanded our design to house 18 strains and have begun taking data with this new chip to provide larger datasets to the classifier. A detailed CAD drawing of this device can be seen in section 6.9 below (Fig. 24). The 18 cell trapping regions are vacuum-loaded from 18 downstream fluidic ports. Long, serpentine channels upstream of the trapping regions (shown in red) serve as fluidic buffer zones to prevent the cross-contamination of strains while loading. The toxin concentration is controlled dynamically by a peristaltic pump that mixes a water/toxin combination (representing
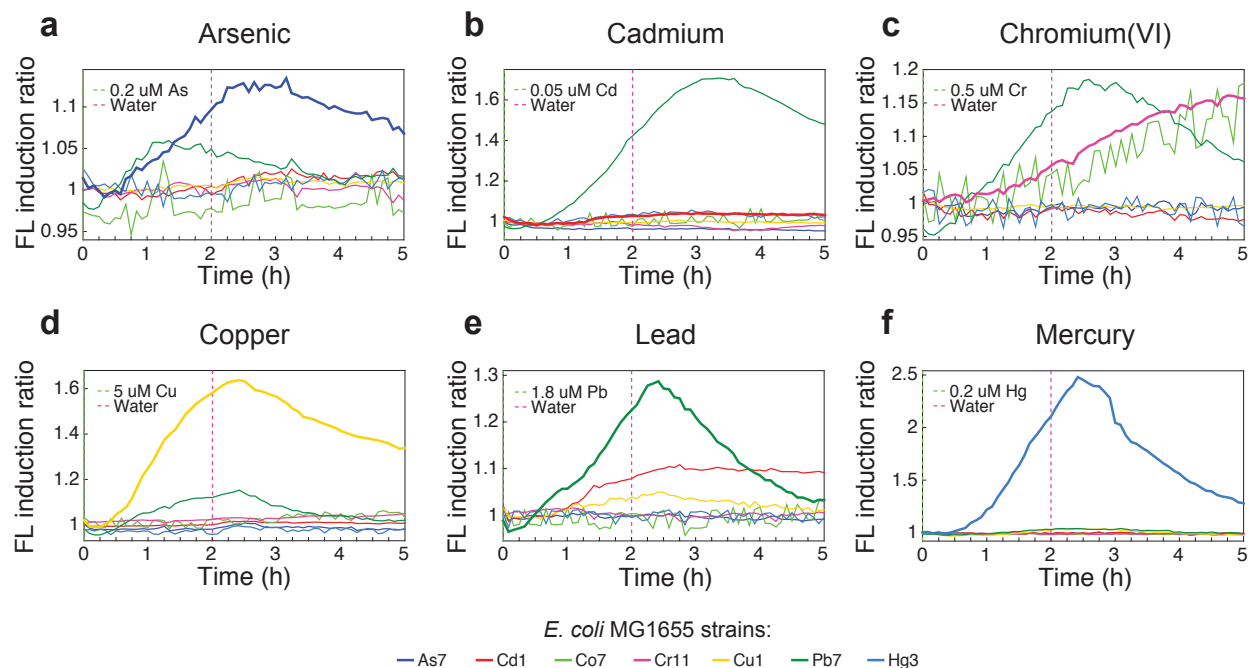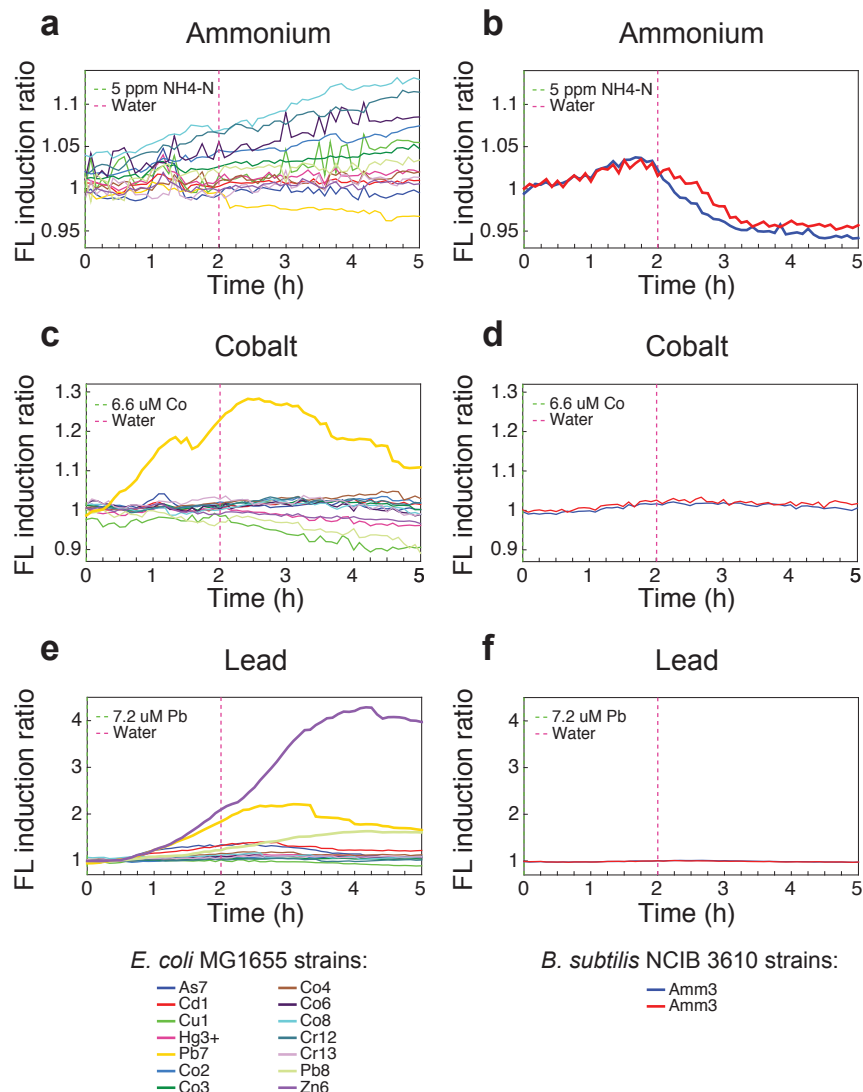
*Figure 2: 8-strain chip induction data for six toxins, demonstrating the sensor-specific induction of the expected strains as well as any response from non-specific strains.*

the source water stream) with a concentrated medium stock.

Importantly, our ability to house many strains on the same chip and subject them to the same toxin inductions enables us to look not only at toxin-specific induction but at crosstalk as well. In Figure 2, we show a single induction time series plot for each of the following six toxins: arsenic, cadmium, chromium(VI), copper, lead, and mercury. Each induction plot shows the fluorescence induction ratio (calculated as (mean fluorescence following induction)/(mean fluorescence prior to induction)) for multiple strains being cultured in the chip, where the toxin is introduced at $t = 0$ and replaced by pure water at $t = 2$ h. In each panel, the time series data for the "on-target" responding strain is plotted in bold. As expected, the GFP fluorescence traces for the "on target" responding strains increase with induction and decrease with un-induction, with some delay due to the cellular response machinery. Figure 2a shows strain *E. coli* MG1655/As7 sensitively responding to 0.2 $\mu$M arsenic, with *E. coli* MG1655/Pb7 exhibiting a smaller response. We have discovered in our experiments that strain *E. coli* MG1655/Pb7 is a less specific responder that aids in toxin identification mainly through combinatorics. Figure 2b shows strain *E. coli* MG1655/Cd1 and *E. coli* MG1655/Hg3 responding to 0.05 $\mu$M cadmium, with *E. coli* MG1655/Pb7 exhibiting a much larger response. Figure 2c shows strain *E. coli* MG1655/Cr11 and *E. coli* MG1655/Co7 responding to 0.5 $\mu$M chromium(VI), with *E. coli* MG1655/Pb7 exhibiting an equal magnitude but more rapid response. Figure 2d shows strain *E. coli* MG1655/Cu1 sensitively responding to 5 $\mu$M copper, with *E. coli* MG1655/Pb7 exhibiting a smaller response. Figure 2e shows strain *E. coli* MG1655/Pb7 responding to 1.8 $\mu$M lead, with *E. coli* MG1655/Cd1 and *E. coli* MG1655/Cu1 responding at lesser magnitudes. Finally, Figure 2f shows strain *E. coli* MG1655/Hg3 sensitively responding to 0.2 $\mu$M mercury.

Within the last two weeks of the reporting period, we were able to collect toxin re-

**Figure 3: 18-strain chip induction data for three toxins, demonstrating the sensor-specific induction of**
**B. subtilis *NCIB 3610/Amm3 for ammonium and the unique combinations of responses allowing the iden-*
*tification of cobalt and lead.***

sponse data across 18 unique sensing strains using a new chip design that houses 16
*E. coli* strains and 2 *B. subtilis* strains. We initially used this chip to test our ammonium
sensor, with a representative induction shown in Figure 3a and 3b. Our *B. subtilis* NCIB
3610/Amm3 construct is a "lights-off" sensor, meaning that GFP production drops when
ammonium is introduced, following a slight increase due to faster growth on this pre-
ferred nitrogen source. We observe this behavior in response to 5 ppm $NH_4$-N in Fig-
ure 3b. The two *B. subtilis* "gill" regions harboring the Amm3 strain respond sensitively
to ammonium, whereas various strains in the *E. coli* "gill" regions shown in Figure 3a
do not. Additional toxin exposures of previously untested strains in this 18-strain chip
demonstrated our ability to sense cobalt, an additional toxin not included in our primary
list of eight. We discovered that strain *E. coli* MG1655/Pb7 responds sensitively to cobalt,
whereas no other *E. coli* or *B. subtilis* strain does (see Fig. 3c, d). We also learned that
strains *E. coli* MG1655/Zn6, *E. coli* MG1655/Pb7, and *E. coli* MG1655/Pb8 respond sen-

6

sitively to lead in descending magnitude, whereas no other strains do (see Fig. 3e, f). By combining these sensor responses, we can identify cobalt by the unique response of Pb7 alone and lead by the unique combination of Zn6, Pb7, and Pb8 responses.

To date, we have performed 247 on-chip toxin exposures with these five prototype sensors, capturing a total of 109,000 transmitted light and fluorescence images. We have run a single microfluidic chip for up to 50 days, which includes inducing and un-inducing the strains with various toxins twice each day. Figure 4 shows the response of a single copper-sensitive strain periodically exposed to toxins and imaged over a period of 34 days, where each row contains approximately three days of data. Mean fluorescence of cells within the trapping region is plotted over time, and toxin exposure events are color-coded. Exposures to "on-target" toxins are shaded red, and exposures to "off-target" toxins are shaded grey. Double-toxin exposures that include the "on-target" toxin are shaded orange. Toxin-spiked water samples were drawn through peristaltic pump tubing to the cells with delays ranging from 1–9 hours, depending on the pump speed and length of tubing. Current long delays are due solely to slow pump rates and long tubing lengths in our initial prototype and not to any fundamental characteristic of the genetic circuits or microfluidic devices. As such, they will be minimized in future design iterations. To visualize and accurately determine the moment of cellular exposure to toxin, on 1/27 we began introducing a 20-minute pulse of green fluorescent tracer dye followed by a 20-minute pulse of pure water before flowing in the toxin. These green fluorescent pulses are shaded green.

In Figure 4, all the "on-target" copper exposures are followed by a rise in mean cellular fluorescence. This indicates that the *E. coli* MG1655/Cu1 strain responds sensitively to copper. Conversely, the majority of "off-target" toxin exposures are not followed by a rise in mean fluorescence. This indicates that the *E. coli* MG1655/Cu1 strain responds quite specifically to copper. For comparison, Figure 5 shows the response of the *E. coli* MG1655/Hg3 strain to various toxin exposures over the same 34-day period. Our initial "on-target" mercury exposures through 2/16 were intended to establish our sensing ability. We subsequently probed the detection limit of this strain by flowing in low concentrations of mercury, which did not elicit a response. Strain *E. coli* MG1655/Hg3 also exhibits less specificity than strain *E. coli* MG1655/Cu1, as indicated by rises in mean fluorescence following "off-target" inductions. This has not been problematic, as we have identified a unique combination of strain responses that is specific for mercury. Movies of the 8-strain single-chip experiment analyzed in Figures 4 and 5 can be found at *http* : //*biodynamics.ucsd.edu*/*DARPA_movies*/*index_DARPA.html*.
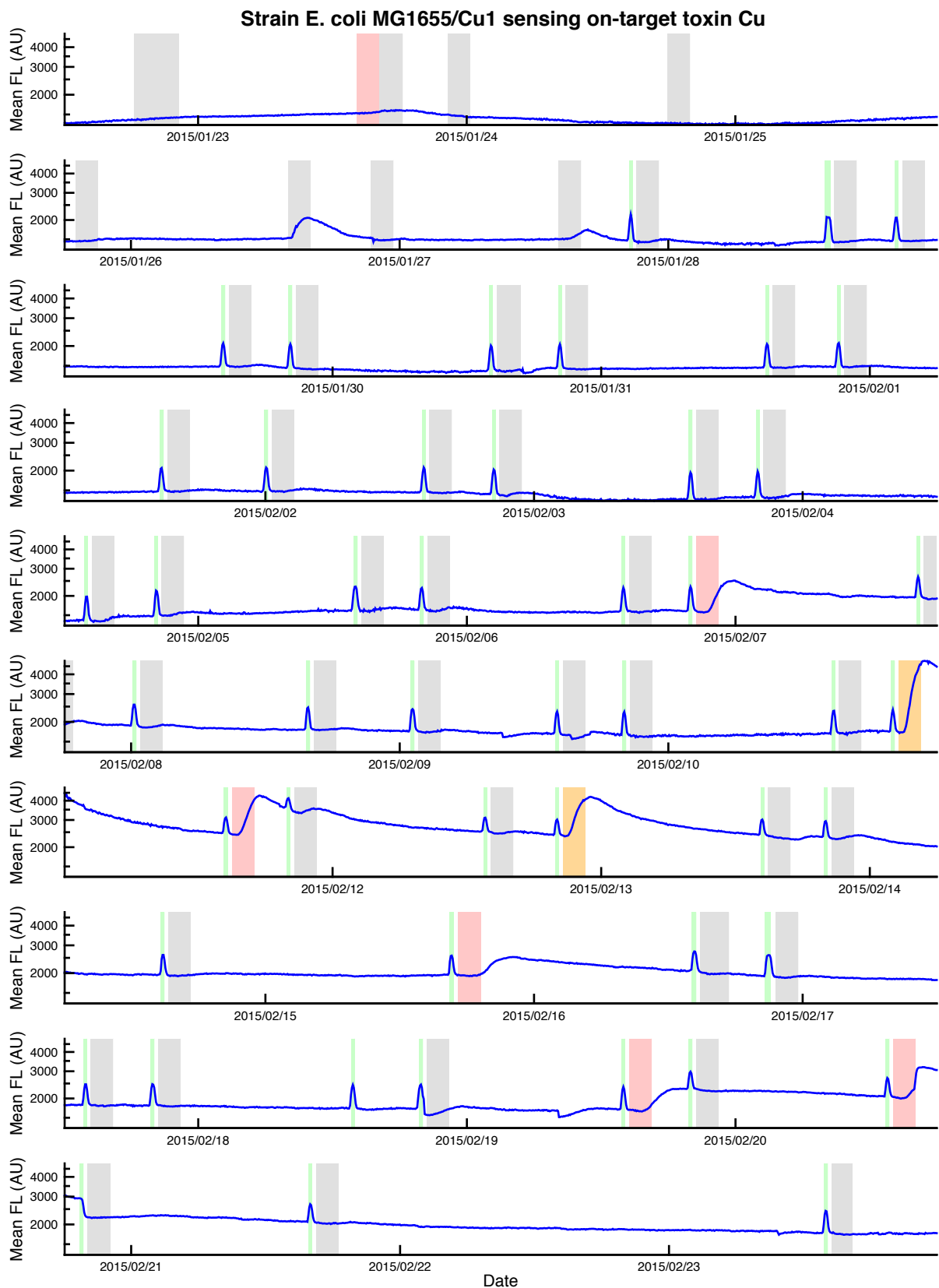
**Figure 4:** *34 days of continuous fluorescence data from the* **E. coli** *MG1655/Cu1 strain as it responds to two inductions per day of various toxins and concentrations.*
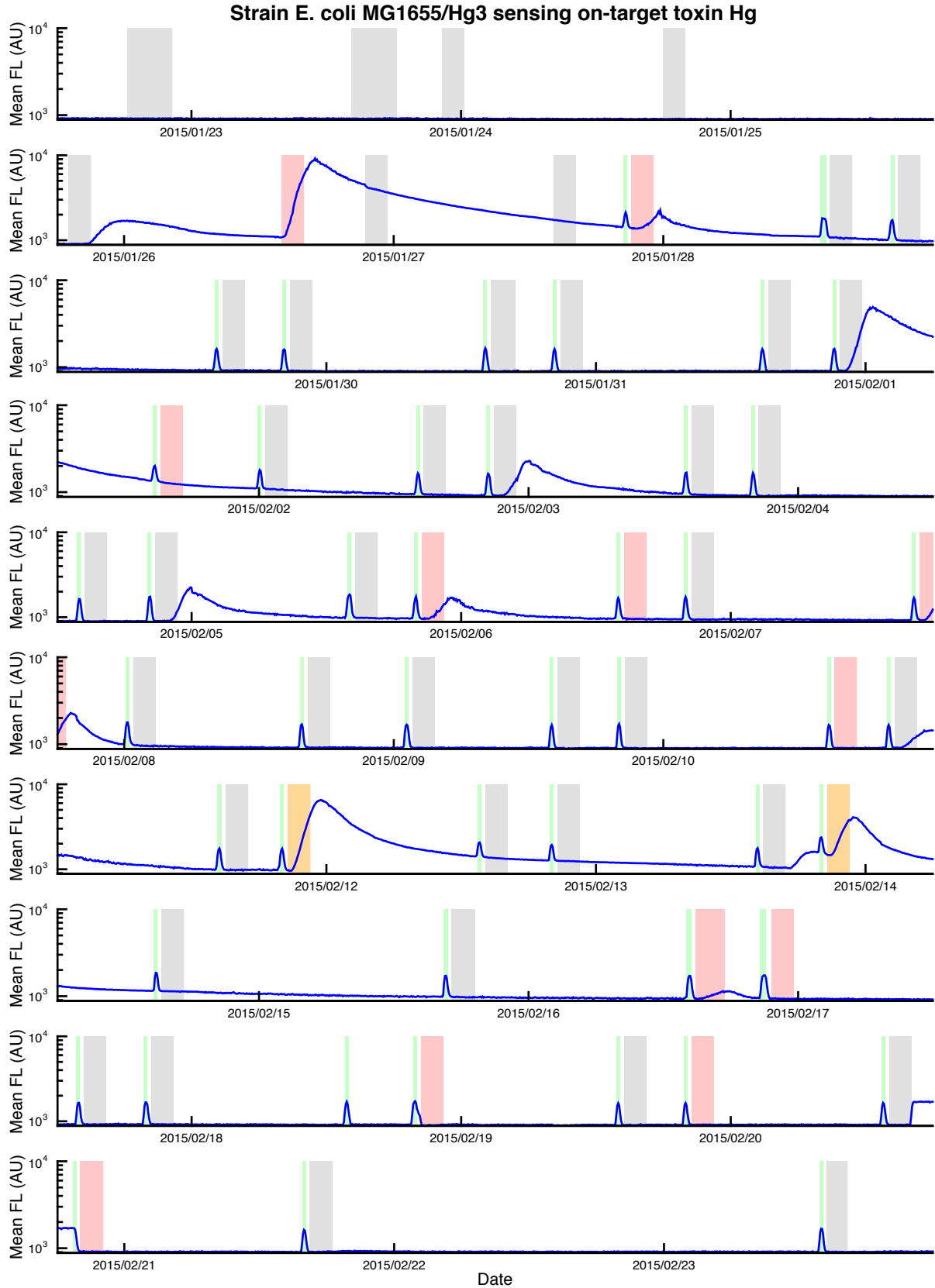
Figure 5: 34 days of continuous fluorescence data from the **E. coli** *MG1655/Hg3 strain as it responds to two inductions per day of various toxins and concentrations.*
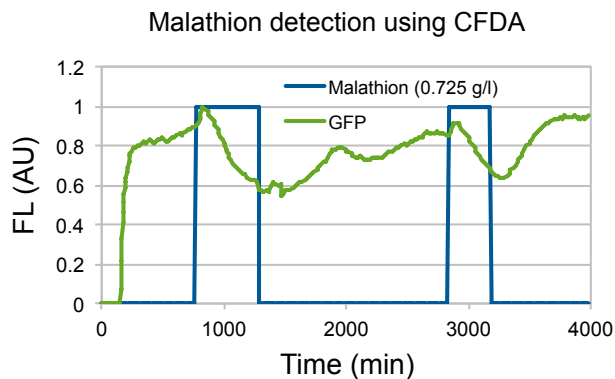
9

**Figure 6: Inhibition of cellular esterase activity in S. cerevisiae during malathion exposure.**

We were successful in transitioning at least one candidate toxin-sensitive promoter for each toxin from RNA-Seq analysis to microscope experiments to prototype sensor experiments with the exception of our initial construct based on the single promoter candidate for malathion (Mal1), which did not produce a discernable response upon loading into the sensor device. This surprised us, and we continued to generate additional constructs based on this promoter while we pressed ahead on collecting on-chip response data for all other sensing strains using the sensor prototypes. After some investigation, we discovered a more robust malathion detection scheme that leverages its esterase inhibition effect in yeast. We loaded *S. cerevisiae* into a gill chip and added CFDA (carboxyfluorescein diacetate) to the growth medium. Because CFDA permeates the cell membrane and is cleaved by cellular esterases into a fluorescent product, we expected the inhibition of esterase activity by malathion to decrease cell fluorescence. This is what we observed during two sequential exposures to malathion in Figure 6. GFP signal significantly drops in the presence of 0.725 g/l malathion, both at around 1,000 and 3,000 min, and subsequently recovers. Because CFDA breaks down in the medium over time thereby increasing the background signal, we introduced fresh medium + CFDA at 1,500 and 2,000 min. This explains the drift in baseline GFP during this period, which can be eliminated in subsequent experiments by mixing dye and medium on-chip.

Our successful measurement of esterase inhibition by malathion in wildtype yeast illustrates the power of our forthcoming whole-genome detection scheme, whereby we will continually measure the responses of 2,000 *E. coli* and *S. cerevisiae* promoters arrayed in a microfluidic chip. We expect the observed esterase inhibition to be significantly represented in the genomic response of this multi-strain library during malathion exposure.

## Task A

**Task Objective:** Identify a combination of cellular signals that will indicate the detection of specific targeted chemical agents

**Milestone:** 12 months

**Metrics/Completion Criteria:**

Milestone 1: Create an initial library of transcription based sensors

Milestone 2: Massively expand the library of transcription based sensors using Next Generation Sequencing (NGS) techniques

**Deliverable:** Comprehensive list of candidate genes in *E. coli* that respond to target compounds

## Reporting Updates

**Milestone 1: Create an initial library of transcription based sensors**

*Summary:* We have completed this milestone. We created a list of toxin responsive promoters based on literature research in Milestone 1.1. We assembled promoters from multiple bacterial species along with necessary regulatory genes into synthetic expression constructs in Milestone 1.2. We obviated the need to intermediately test the response of the expression constructs in batch culture using a plate reader in Milestone 1.3 by directly validating within custom microfluidic devices in Milestone 1.4.

*Milestone 1.1:* *Identify 3 promoters for each toxin of interest based on literature searches:* We have identified 3 or more promoters from literature for each toxin, when available. The exceptions are copper (2), lead (1), and malathion (0), for which less than three promoters are known.

Bacteria that grow in environments with high toxin levels, occurring either naturally or as a result of pollution, have evolved pathways to mitigate their effects on cellular metabolism. These resistance pathways are often activated by a specific transcription factor that is sensitive to the intracellular concentration of the toxin. We searched the scientific literature to identify well-characterized transcription factor/promoter pairs. Where divergent pathways responding to the same toxin have been characterized, we chose at least one representative pathway from each major evolutionary clade. Identified candidate regulatory mechanisms for specific toxins, including the native organism and naturally-occurring DNA construct, are provided below.

*Arsenic:* The arsR family of transcription factors contains many arsenic-sensing members, including one found on the *E. coli* genome[1] and another encoded on the *E. coli* R773 plasmid.[2] We selected these two candidates for synthesis because they are native to *E. coli* and have been used previously to construct a biosensor.[3] We also synthesized an arsR construct from *S. aureus*.

*Cadmium:* Cadmium-responsive transcription factors have been identified from both the arsR family[4] and the merR family[5] of transcriptional regulators. We selected a representative member from each group: cadC from *S. aureus*[6] and cadR from *P. putida*.[7]

*Chromium(VI):* We selected the chromate-responsive element from *O. tritici*, which has been characterized[8] and used to construct a highly specific biosensor,[9] a related system from *C. metallidurans*, and an unrelated chromate-responsive transcription factor from the genome of *B. subtilis* NCIB 3610.[10]

*Copper:* We selected two well-studied functionally-unique copper-responsive elements native to *E. coli*.[11] CueR functions as an activator/repressor of the merR family, while cusS/R is a two-component sensor system including a histidine kinase, which

may aid in signal amplification.

*Lead:*   The only lead-specific system identified in the literature is pbrR,[12] found on a megaplasmid in *C. metallidurans*.[13]

*Mercury:*   Mercury-inducible merR systems have previously been used to construct biosensors with low detection limits,[14] where merR functions as a repressor in the absence of mercury and an activator in the presence of mercury. We adapt the same strategy for our biosensor plasmids, selecting three well-studied members of the merR family.[15,16,17]   Another more evolutionarily divergent merR protein from *S. lividans* was not pursued because it was shown to function only as a repressor.[18]

*Ammonium:*   Sensing ammonium requires culturing cells in a background nitrogen source that is less-preferred than ammonium. While nitrate is the most suitable background nitrogen source for this purpose due to its long-term stability in solution, *E. coli* MG1655 is unable to assimilate it under aerobic conditions. Therefore, we decided to sense ammonium using the bacterium *B. subtilis* NCIB 3610, which can utilize both nitrate and ammonium as a nitrogen source. We incorporated the native ammonium-sensitive promoters $\mathbf{p}_{nasA}$ and $\mathbf{p}_{nasB}$ and the synthetic promoter $\mathbf{p}_{spo1-tnrA}$ into plasmids for integration into a neutral site on the *B. subtilis* NCIB 3610 genome, where they act as a second copy of the promoter to drive GFP expression in the presence of ammonium.

For completeness, in Table 1 we present the full list of identified candidate constructs. Sensor plasmids containing the most promising regulatory candidates (shown in boldface in Table 1) for arsenic, cadmium, chromium(VI), copper, lead, and mercury were synthesized by our cloning vendor and ported into *E. coli* to drive GFP expression in the presence of the toxin. Synthesized sensing constructs for ammonium were integrated into a modified version of

| Toxin | Source Organism | Gene/Promoter |
|---|---|---|
| Arsenic | ***E. coli* plasmid** | $\mathbf{\mathit{arsR}/p_{arsR}}$ |
| | ***E. coli* genome** | $\mathbf{\mathit{arsR}/p_{arsR}}$ |
| | ***S. aureus* plasmid** | $\mathbf{\mathit{arsR}/p_{arsR}}$ |
| | *B. subtilis* genome | $\mathit{arsR}/p_{arsR}$ |
| Cadmium | ***S. aureus* plasmid** | $\mathbf{\mathit{cadC}/p_{cadC}}$ |
| | ***P. putida* genome** | $\mathbf{\mathit{cadR}/p_{cadR}}$ |
| | *S. salivarius* genome | $\mathit{cadX}/p_{cadX}$ |
| | *S. lugdunensis* genome | $\mathit{cadX}/p_{cadX}$ |
| Chromium(VI) | ***C. metallidurans* plasmid** | $\mathbf{\mathit{chrB}/p_{chrB}}$ |
| | ***O. tritici* transposon** | $\mathbf{\mathit{chrB}/p_{chrB}}$ |
| | ***B. subtilis* genome** | $\mathbf{\mathit{chrS}/p_{chrS}}$ |
| Copper | ***E. coli* genome** | $\mathbf{\mathit{cueR}/p_{copA}}$ |
| | ***E. coli* genome** | $\mathbf{(\mathit{cusS/R})/p_{cusC}}$ |
| Lead | ***C. metallidurans* plasmid** | $\mathbf{\mathit{pbrR}/p_{pbrR}}$ |
| Mercury | ***E. coli* plasmid** | $\mathbf{\mathit{merR}/p_{merR}}$ |
| | ***S. aureus* plasmid** | $\mathbf{\mathit{merR}/p_{merR}}$ |
| | ***S. marcescens* plasmid** | $\mathbf{\mathit{merR}/p_{merR}}$ |
| | *S. lividans* genome | $\mathit{merR}/p_{merR}$ |
| Ammonium | ***B. subtilis* genome** | $\mathbf{p_{nasA}}$ |
| | ***B. subtilis* genome** | $\mathbf{p_{nasB}}$ |
| | ***B. subtilis* genome** | $\mathbf{p_{spo1-tnrA}}$ |

*Table 1: Identified promoters from the literature.*

the *B. subtilis* NCIB 3610 host strain, where a motility gene (*hag*) and a biofilm pathway gene (*epsH*) were knocked out for improved growth within microfluidic devices.

***Milestone 1.2:***   *Clone the promoters into an expression plasmid driving the production of GFP:* Completed in Q1. We have cloned all promoters identified from the literature into a standardized plasmid expression system.

We synthesized two standardized plasmid backbones for testing the candidate toxin-responsive elements identified in the literature in *E. coli*. Both vectors include the p15A medium-copy origin of replication, a spectinomycin resistance cassette for selection of positive transformants, and a promoter-less GFP insulated by flanking terminators. One of the vectors replaces the native ribosome binding site (RBS) in front of GFP with a version known to produce high levels of expression in *E. coli* MG1655.

Both synthesis of the plasmid backbones and insertion of the sequences for candidate toxin-responsive elements into each version were carried out by Transcriptic. The sequence of interest was inserted such that the transcriptional regulator remains under the control of its native promoter, with the inducible promoter driving expression of GFP. Because all DNA constructs were completely synthesized, we were able to codon optimize sequences for improved heterologous expression in *E. coli* MG1655.

***Milestone 1.3:*** *Validate the response of the promoters using traditional batch experiments (Note validation indicates a ratio of induced fluorescence signal to uninduced of at least 3:1):* Completed in Q1. Instead of performing preliminary testing in batch, we validated the response of the promoters directly in our microfluidic devices.

We expected the development of our microfluidic devices to lag the construction of our toxin-sensing plasmids, thereby requiring that initial induction experiments be performed using a fluorescent plate reader. However, rapid development of the microfluidics allowed us to run initial induction experiments of the sensing strains on-chip, and, upon successfully inducing sensing strains on-chip, we saw no need to replicate experiments in batch culture for several reasons. First, since cellular measurements are highly dependent upon growth phase, and batch culture cannot provide a constant growth environment, we expect more reproducible results from our microfluidic chemostats. Second, inducing on-chip allows us to dynamically control the inducer concentration and thereby more efficiently scan the induction range for each sensing strain. Finally, our deployable biosensor device will incorporate microfluidic culturing and optical imaging; therefore, microscope imaging on-chip is more relevant than reading fluorescence values within the batch culture wells of a plate reader.

***Milestone 1.4:*** *Validate the response of the promoters using microfluidic devices. Quantitatively measure GFP signal in response to various relevant levels of toxins of interest:* We have validated the response of the promoters in microfluidic devices for all toxins of interest.

In Q1, we developed a microfluidic device to culture and induce each sensing strain with varying levels of toxins of interest. Termed the "gill chip," this device is a variation on the "biopixel" device previously developed in our lab. Structurally, this microfluidic device consists of a polydimethylsiloxane (PDMS) elastomer block with recessed channels that are sealed upon bonding to a glass cover slip. Fluidically, the core functional unit of the device is a microfluidic channel with long, narrow branches along the sides that serve to retain cultures of cells (Fig. 7). The precise geometry of the "cell trap" creates high-resistance, low-flow regions where cells grow in a defined area with continuous perfusion of fresh medium. Previous studies of synthetic gene circuits required microfluidic devices which grow cells in a monolayer in order to perform single-cell tracking. The gill chip uses a similar side trap geometry, but the trap width is decreased from 100 $\mu$m to 10 $\mu$m, and the trap height is increased from 1.65 $\mu$m to 50 $\mu$m. By using thin, tall traps, the optical signal is increased while maintaining the ability to selectively grow cells in a low-flow trap region. The overall device consists of an array of trapping units connected in parallel, each supporting the long-term culturing of a cell strain. The central channels of the trapping regions are fed by a single switchable medium source.

We optimized the length of the cell-trapping channels of the gill chip to maximize fluorescent signal while ensuring adequate medium delivery for healthy cell growth. Medium delivery to the packed cells in the trapping channels is limited
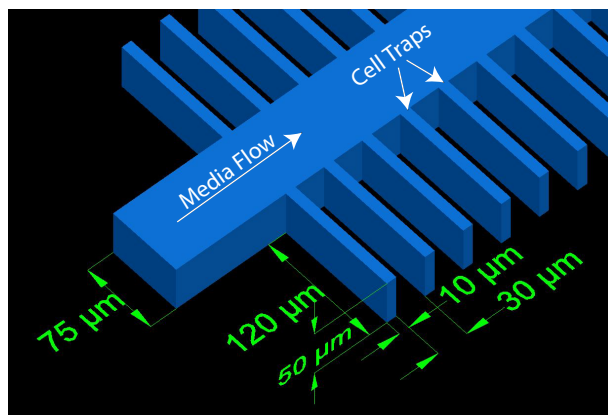
13

Figure 7: Functional unit of the microfluidic "gill chip." Medium is delivered via the main channel, and cells are cultured in the branched channels.



Figure 8: Optimization of cell trap length based on GFP expression level. 120-$\mu$m-long traps provide the greatest fluorescent signal for all arsenic concentrations.

by diffusion, and we observed that excessively long channels slowed cell growth and GFP production. We determined the optimal trap length by fabricating multiple versions in parallel and measuring expressed fluorescence from a preliminary arsenic-sensing strain. In Figure 8, we see that a 120 $\mu$m channel length produces the largest fluorescent signal for arsenic levels ranging from below 1$\times$ to $\approx$10$\times$ the EPA limit. Therefore, we standardized our trap length in future experiments at 120 $\mu$m. We confirmed that representative gram-negative (*E. coli* MG1655, *S.* Typhimurium) and gram-positive (*B. subtilis* NCIB 3610) bacterial strains as well as yeast (*S. cerevisiae* MFSC120) form stable cultures in the gill chip with 120-$\mu$m-long trapping channels. Initial experiments continuously growing *E. coli* MG1655 in LB medium in the gill chip demonstrated a device lifetime of 27 days before flow stopped due to channel clogging. This lifetime was later extended to **50 days** by slowing the cellular growth rate by replacing LB medium with M9 minimal medium.

While waiting for Transcriptic to synthesize our initial library of toxin-sensitive plasmids based on our literature search, we used our microfluidic device to measure the
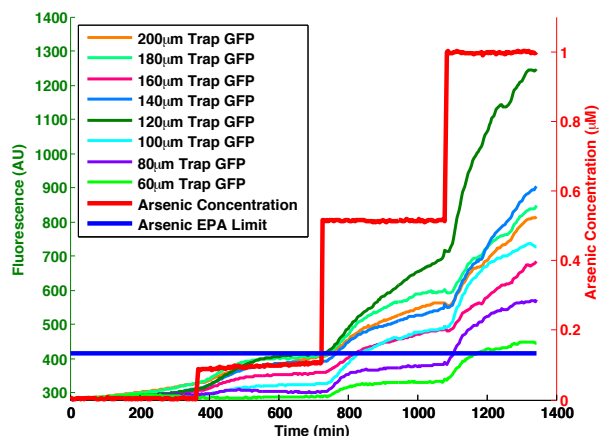
response of two toxin-sensing promoters in *E. coli* MG1655 generated in our lab for previous work.

We first tested induction of pRS18, an arsenic-sensing plasmid, in our microfluidic device. A recombinant strain of *E. coli* MG1655 containing pRS18 was vacuum-loaded into the microfluidic device, grown to confluence, and step-induced with sodium ortho-arsenite ($Na_3AsO_3$) in M9 minimal medium supplemented with 0.4% glucose. Images were collected using a Nikon Ti microscope with 4$\times$ objective magnification. Step induction was executed using our gravity-driven "Dial-A-Wave" automated flow control system to mix two medium sources in a defined ratio. The cells were sequentially exposed to arsenic concentrations of 0 $\mu$M, 0.1 $\mu$M, 0.5 $\mu$M, and 1 $\mu$M in 6-h windows, allowing time for the cells to respond to each concentration increase. The data shows a significant detection of 0.1 $\mu$M arsenic, which is below the EPA limit of 0.13 $\mu$M. Figure 9 shows the trapped cells prior to (panel a) and following (panel b) 6 h of induction at 0.1 $\mu$M arsenic. We believe that the rise in GFP expression prior to the induction step is due to a small leak of inducer into
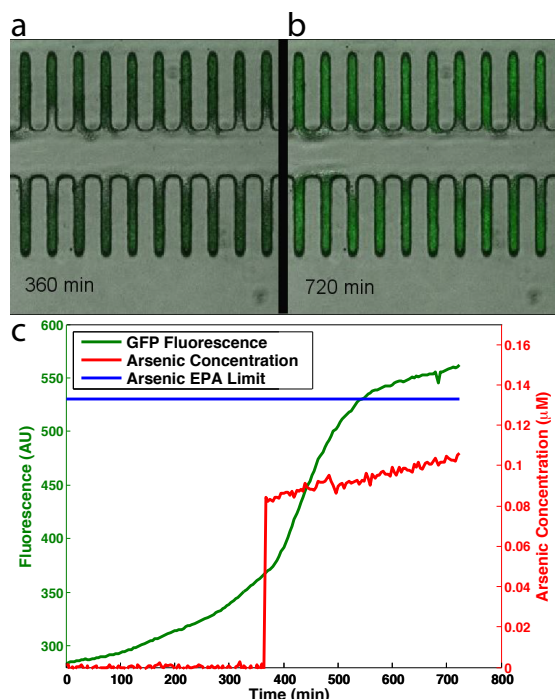
*Figure 9: Response of arsenic-sensing plasmid pRS18 (from previous work) to step-induction with arsenic. Images before (a) and after (b) induction with 0.1 μM arsenic. c) Measured GFP expression over time for induction with 0.1 μM arsenic.*

*Figure 10: Response of copper-sensing plasmid pCueCopA (from previous work) to step-induction with copper. Measured GFP expression over time for induction with 25 μM copper demonstrates that we can detect copper concentrations near the EPA limit.*

the medium stream. Average fluorescence within the trapping regions of the images as well as arsenic concentration is plotted over time in Figure 9c.

Second, we tested induction of pCue-CopA, a copper-sensing plasmid, in our microfluidic device. A recombinant strain of *E. coli* MG1655 containing pCueCopA was vacuum-loaded into a microfluidic device, grown to confluence, and step-induced with copper sulfate ($CuSO_4$) in LB medium. Images were collected using an Etaluma LumaScope microscope with $20\times$ objective magnification. Step induction was again performed using our "Dial-A-Wave" automated flow control system. The cells were sequentially exposed to copper concentrations of 0 μM, 25 μM, 52 μM, 77 μM, and 99 μM in 6-h windows. Figure 10 shows a significant detection of 25 μM copper, which is slightly above the EPA limit of 20.5 μM.

In Q2, we received our literature-based library of synthesized sensor plasmids from Transcriptic and transformed them into both *E. coli* MG1655 and *E. coli* LABEC01 (See Detailed Methods for strain details) for on-chip testing. To increase the throughput of our microfluidic experiments, we developed a version of the gill chip with four independently loadable cell growth areas for culturing four sensing strains in parallel (see Fig. 11). These growth areas are fed from the output of an on-chip "Dial-A-Wave" system to dynamically control the local toxin concentration.

For on-chip induction experiments, strains were loaded into the device and grown in M9 minimal medium with 0.4% glucose and 50 μg/ml spectinomycin (for plasmid maintenance) for 2 days in the absence of toxin. Cells were imaged under $4\times$ brightfield and GFP fluorescence filters on an Olympus IX-81 microscope every 2 min for 6 h under these non-inducing conditions to generate a fluorescence baseline. We then introduced the toxin to induce GFP expression and imaged the cells every 2 min for 6 h.

15
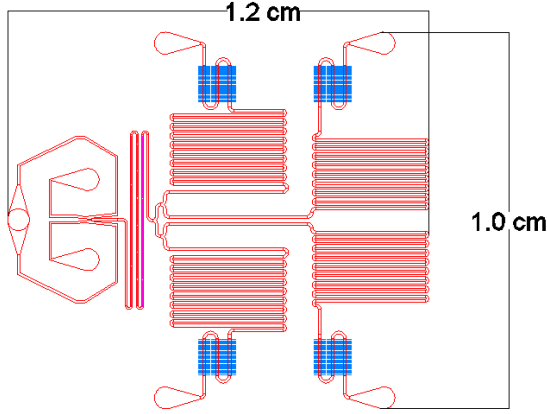
*Figure 11: Gill chip v9-C for parallelizing the on-chip testing of toxin-responsive promoters. Four cell trapping regions (shown in blue) are vacuum-loaded from four downstream fluidic ports. Long, serpentine channels upstream of the trapping regions (shown in red) serve as fluidic "buffer zones" to prevent the cross-contamination of strains. The three-port "Dial-A-Wave" module at the left serves to mix two medium streams to precisely and dynamically define the toxin concentration in the cell trapping regions.*

To determine the response to toxin induction, we compared the raw fluorescence images before and after induction. The image stacks were analyzed using a custom ImageJ script to measure the average GFP signal over time for each cell trap. Any linear trend in the uninduced fluorescence data due to cell growth was subtracted from the data. In this manner, the dynamic signal-to-noise ratio (SNR) was calculated as: (current fluorescence - mean uninduced fluorescence)/(standard deviation of uninduced fluorescence).

We used this microfluidic platform to measure the response of several toxin-responsive promoter constructs identified from the literature and synthesized by Transcriptic. Table 2 summarizes the results of these experiments, listing the most sensitive promoter constructs exposed to each toxin and the calculated SNR values after 6 h of induction. Note that SNR values greater
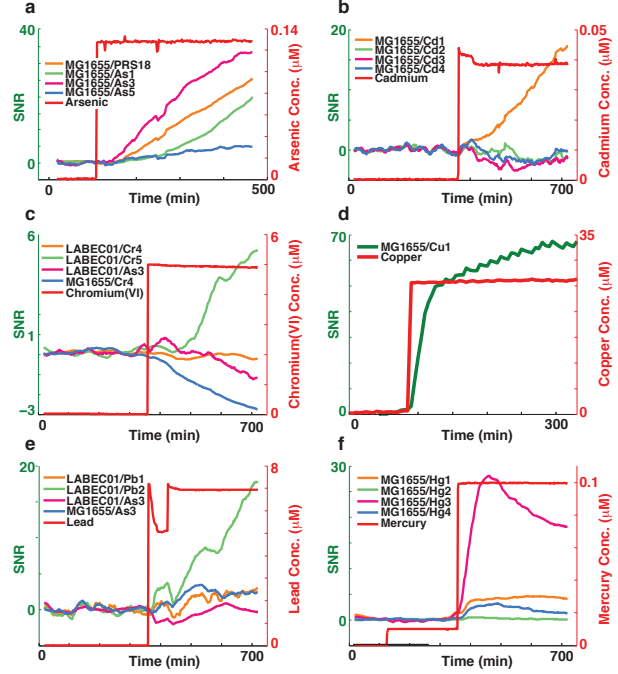


*Figure 12: On-chip time-lapse induction responses of promoter constructs identified from the literature, expressed in E. coli, and summarized in Table 2. a) MG1655/As3, MG1655/pRS18, and MG1655/As1 show high sensitivity to 0.13 μM arsenic. (Note that pRS18, a plasmid generated in our lab for previous work, shares the same promoter as Transcriptic plasmid As7.) b) MG1655/Cd1 shows high sensitivity to 0.04 μM cadmium. c) LABEC01/Cr5 shows high sensitivity to 5 μM chromium(VI). d) MG1655/Cu1 shows high sensitivity to 25 μM copper. e) LABEC01/Pb2 shows high sensitivity to 7 μM lead. f) MG1655/Hg3 shows high sensitivity to 0.1 μM mercury.*

than ≈2 indicate a significant response.

The on-chip time-lapse induction response of each promoter construct in this table can be seen in Figure 12. In each figure panel, the moment of toxin introduction is represented as a step in the red trace. While some promoter constructs exhibit a sensitive response by highly expressing GFP immediately following induction, others do not. Note that the arsenic sensor plasmid As3 was used as a negative control for experiments with non-arsenic-sensing strains, and it showed no induction in the presence of other heavy metals.

16

| Toxin | Gene/ Promoter | Source | RBS | Host Strain/ Plasmid | Concentration Sensed in Microfluidic Device ($\mu$M) | SNR after 6 h | Refer to Figure Panel |
|---|---|---|---|---|---|---|---|
| Arsenic | $arsR$/$p_{arsR}$ | *E. coli* plasmid | native | *E. coli* MG1655/As1 | 0.13 | 20 | 12a |
| Arsenic | $arsR$/$p_{arsR}$ | *E. coli* genome | synthetic | *E. coli* MG1655/As3 | 0.13 | 33 | 12a |
| Cadmium | $cadC$/$p_{cadC}$ | *S. aureus* plasmid | native | *E. coli* MG1655/Cd1 | 0.04 | 17 | 12b |
| Chromium(VI) | $chrB$/$p_{chrB}$ | *O. tritici* transposon | native | *E. coli* LABEC01/Cr5 | 5 | 5 | 12c |
| Copper | $cueR$/$p_{copA}$ | *E. coli* genome | native | *E. coli* MG1655/Cu1 | 25 | 65 | 12d |
| Lead | $pbrR$/$p_{pbrR}$ | *C. metallidurans* plasmid | synthetic | *E. coli* LABEC01/Pb2 | 7 | 18 | 12e |
| Mercury | $merR$/$p_{merR}$ | *E. coli* plasmid | synthetic | *E. coli* MG1655/Hg3 | 0.1 | 20 | 12f |

*Table 2: Toxin-responsive promoter constructs identified from the literature, synthesized by Transcriptic, cloned into* **E. coli**, *and demonstrating high sensitivity in the microfluidic device, ordered by toxin. The promoter source and RBS used in the synthetic construct are shown alongside the toxin concentration sensed and SNR after 6 h. Refer to Figure 12 for on-chip time-lapse induction responses.*

| Toxin | Conc ($\mu$M) | Strain/On-Target Toxin | | | | | |
|---|---|---|---|---|---|---|---|
| | | As3/ Arsenic | Cd1/ Cadmium | Cr5/ Chromium(VI) | Cu1/ Copper | Pb2/ Lead | Hg3/ Mercury |
| Arsenic | 1.3 | 1.00 | -0.18 | 0.24 | -0.05 | 0.33 | 0.25 |
| Cadmium | 0.44 | -0.13 | 1.00 | -0.02 | -0.08 | 0.46 | -0.26 |
| Chromium(VI) | 2 | -0.22 | 0.62 | 1.00 | -0.06 | 1.69 | -0.96 |
| Copper | 20 | | -0.09 | -0.08 | 1.00 | -0.17 | 0.16 |
| Lead | 1 | -0.05 | 0.06 | 0.01 | -0.05 | 1.00 | 0.12 |
| Mercury | 0.1 | -0.18 | -0.09 | 0.58 | -0.05 | 1.44 | 1.00 |
| Cobalt | 20 | 0.30 | 0.67 | -0.11 | -0.01 | 0.80 | |
| Iron | 500 | -0.06 | -0.31 | 0.86 | 0.17 | 0.96 | -0.38 |
| Nickel | 1.7 | -0.08 | -0.08 | -0.35 | -0.05 | 0.16 | |
| Zinc | 7.6 | -0.14 | -0.09 | 0.35 | -0.08 | 1.06 | 0.78 |

*Table 3: Specificity results for toxin-sensing plasmids identified from the literature and cloned into synthetic constructs expressed in* **E. coli**. *Here, strains representing the most sensitive promoters from our microfluidic induction experiments were grown in the wells of a microplate in the presence of on- and off-target heavy metals. Each column represents the fluorescence response of a strain to the on-target metal (normalized to "1") and all off-target metals, where "0" represents the unchanged response for the no-toxin control. Although these promoters are generally specific to the on-target toxin, some crosstalk is evident (**i.e.** Cr5 is sensitive to iron; Hg3 is sensitive to chromium(VI) and zinc; Pb2 is sensitive to several off-target toxins).*

To investigate the specificity of the most sensitive promoter construct for each toxin identified from the literature, cloned into synthetic constructs expressed in *E. coli*, and measured on-chip (see Table 2), we grew these strains in the wells of a microplate in the presence of on- and off-target heavy metals. Although the ability to sense cobalt, iron, nickel, and zinc is of secondary importance, we included these heavy metals as potential off-target inducers because they may be present in the natural water supply under test. Each column of Table 3 shows the fluorescence response of a strain to the on-target metal (normalized to "1") and all off-target metals, where "0" represents the unchanged response for the no-toxin control. Although these promoters are generally specific to the on-target toxin, some crosstalk is evident. Fortunately, in cases of significant crosstalk (*e.g.* Pb2 lead-sensing strain), nonspecific promoter responses can be combinatorially combined with other promoter responses to generate specific multi-promoter responses (see Fig. 15 and Appendix B). This combinatorial logic was implemented in algorithms developed in Q3–Q4. Additionally, our literature-identified promoters were supplemented with highly specific promoters via RNA-Seq analysis (*e.g.* ybiI lead-responsive promoter).

**Milestone 2: Massively expand the library of transcription based sensors using Next Generation Sequencing (NGS) techniques**

*Summary:* We have completed this milestone. We expanded upon our proposed work by exposing multiple strains of bacteria (in addition to *E. coli*) to relevant levels of all toxins and extracted RNA in Milestone 2.1. To ensure that the biosensor will have high selectivity for the toxins of interest, we exposed the sensing bacteria to multiple toxins at once and extracted RNA

in Milestone 2.2. We obviated the need for testing at multiple temperatures by ensuring tight temperature control in our prototype enclosure in Milestone 2.3. We optimized our microfluidic devices to achieve growth rates similar to batch culture in Milestone 2.4. We prepared isolated RNA and generated sequence data in Milestone 2.5. We analyzed the sequencing data for all toxins and discovered numerous differentially expressed genes in Milestone 2.6.

***Milestone 2.1:*** *Expose target cells to toxins of interest over four orders of magnitude of concentration and extract RNA:* Completed in Q1. We have exposed three strains of bacteria to relevant levels of all toxins and extracted RNA.

Our toxin exposure protocol involves first growing a culture of bacterial cells to mid-log phase (OD$\leq$0.2). Second, we dilute with fresh media, add the toxin of interest, and culture at 37$°$C for 3 h, ensuring that the cells do not exit exponential growth phase before harvesting (OD$\leq$0.25). Finally, we stabilize cellular RNA using Qiagen RNA Protect reagent, centrifuge the cells, and freeze the pellet at -80$°$C.

To investigate the cellular RNA response to toxin insult, we exposed *E. coli* MG1655, *E. coli* LABEC01, and *B. subtilis* 168 cells to the toxins and concentrations shown in Table 4 and extracted RNA. In determining the toxin concentrations for exposure, we began with the EPA limit. We generated growth curve data for *E. coli* MG1655 cells exposed to each toxin at levels around the EPA limit (Fig. 13) and, combined with our induction data, found that the concentration range of interest between the detection limit and cell death is generally less than 4 orders of magnitude. Surprisingly, toxin levels in this range tend not to slow down growth rate but instead delay the onset of the exponential growth phase, as illustrated by copper exposure in Figure 13a. Figure 13b shows that the dependence of
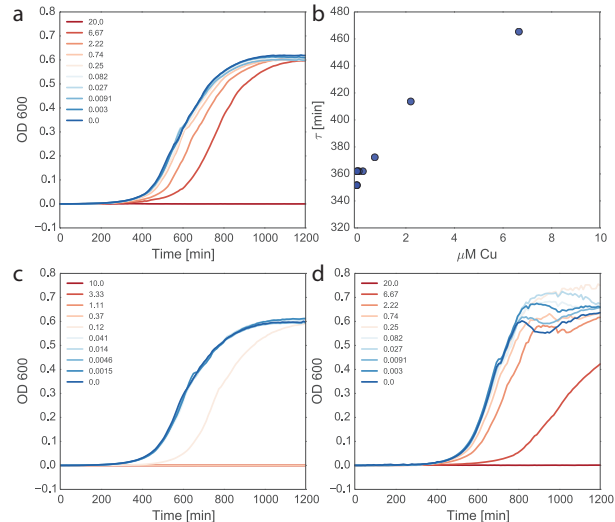
| Toxin | P/S | Exposure Conc. ($\mu$M) | | |
| --- | --- | --- | --- | --- |
| | | *E. coli* MG1655 | *E. coli* LABEC01 | *B. subtilis* 168 |
| **Ammonium** | n/a | | | **71.4, 714** |
| **Arsenic** | P | **0.25, 1** | **0.1, 1** | |
| **Cadmium** | P | **0.4, 1.2** | **0.04, 0.4** | |
| **Chromium(VI)** | P | **0.2, 2** | **0.2, 2** | |
| Cobalt | S | 20, 200 | | |
| **Copper** | P | **0.2, 2, 20** | **0.2, 2** | |
| Iron | S | 5.4, 54 | | |
| **Lead** | P | **0.1, 3** | **0.1, 3** | |
| **Malathion** | n/a | **21.9, 219** | **21.9, 219** | |
| **Mercury** | P | **0.01, 0.1** | **0.01, 0.1** | |
| Nickel | n/a | 0.77, 7.7 | | |
| Zinc | S | 76, 760 | | |

*Table 4: Concentrations used for single-toxin exposures of bacterial strains to investigate the cellular RNA response for sensitivity. Most water toxins to be sensed by our device (boldfaced) are identified by the EPA as being of primary (P) importance. Sensing is expected to occur in the presence of interfering substances of secondary (S) importance (not boldfaced).*



*Figure 13: Batch growth rate data for **E. coli** MG1655 cells exposed to a) copper, c) mercury, and d) chromium(VI). Panel b) shows the approximately linear scaling of lag prior to exponential growth phase with copper concentration.*

growth phase lag on copper concentration is approximately linear. If cells exposed to a toxin at the EPA limit did not exhibit this growth phase lag (see growth rates for mercury exposure in 13c), we assumed that they were not at risk of death and probed additional toxin concentrations that were higher. If cells exposed to a toxin at the EPA limit exhibited this growth phase lag (see growth rates for chromium(VI) exposure in 13d), we assumed that they were at risk of death and probed additional toxin concentrations that were lower. Our general strategy was to examine the cellular RNA response at toxin levels where the cells are responding specifically to the presence of the toxin without triggering a universal (and unspecific) cellular stress response.

***Milestone 2.2:*** *Expose target cells to multiple toxins at once and extract RNA to determine selectivity of sensor response and use computational algorithms to determine a highly accurate relationship between cellular signals and sensor*

*response:* We have exposed cells to multiple toxins at once and determined that the promoters we identified respond as expected, even in the presence of an additional toxin.

We have performed six RNA-Seq experiments with As+Cd, As+Hg, and Cd+Hg in *E. coli* MG1655 and LABEC01 to investigate the specificity of the expression responses. Concentrations used for each multiple-toxin exposure are shown in Table 5. For this subset of toxin combinations, only those RNA-Seq promoter candidates responding to each individual toxin also respond to the same toxin in combination with others (See the rightmost three conditions in Fig. 14 and additional figures in Appendix B). To avoid the complexity of using this RNA-Seq approach combinatorially to investigate specificity across all toxin combinations, we concluded that the most efficient approach is to develop sensor strains based on RNA data extracted following single-toxin exposures and then expose these strains to other toxins to look for non-specific responses.

Because we have several strains that re-

| Toxin Combination | Exposure Conc. ($\mu$M) | |
| --- | --- | --- |
| | *E. coli* MG1655 | *E. coli* LABEC01 |
| Arsenic/Cadmium | 0.25/0.4 | 0.1/0.04 |
| Arsenic/Mercury | 0.25/0.1 | 0.1/0.01 |
| Cadmium/Mercury | 0.4/0.1 | 0.04/0.01 |

*Table 5: Concentrations used for multiple-toxin exposures of* **E. coli** *strains to investigate the specificity of the cellular RNA response.*

spond to each toxin, even for some overlap in response between certain promoters and multiple toxins, we can use computational analysis to determine a disjoint set of genes, and therefore a unique response signature, for each toxin of interest. In Q3–Q4 we exposed the prototype device containing an array of sensing strains to each toxin and used the GFP "fingerprint" (*i.e.* which strains are responding at what intensities) to train a robust computational algorithm to decode this fingerprint into the toxin and concentration present in the sample.

***Milestone 2.3:*** *Repeat above experiments at 25°C, 30°C, and 37°C:* Completed in Q1. All biosensing organisms grow optimally at 37°C. Upon investigating the ability to maintain a constant on-chip temperature (Task C), we have elected to design the sensor with temperature control, as this implementation will be simple and will eliminate the risk of any temperature-related sensitivity or specificity issues.

***Milestone 2.4:*** *Determine how growth system affects gene expression - compare batch grown cells to microfluidic grown and ensure proper growth and response of cells to toxins in device environment:* Completed in Q1. We have developed the microfluidic device and confirmed that the cells in the chambers are healthy, expressing normally, and their growth rates are comparable to those observed in batch conditions.

A simple method to indicate healthy cell growth on-chip is to compare estimated microfluidic growth rate with calculated batch growth rate. We grew *E. coli* MG1655 cells in M9 minimal medium supplemented with 0.4% glucose in the plate reader and extracted a doubling time of 0.93 h from measurements of $OD_{600}$ over time. We grew cells in the microfluidic device with the same medium and estimated the growth rate by collecting the effluent exiting the device, serially diluting it, and plating it on LB medium in agar to determine the viable cell count. Using this data, we determined an on-chip doubling time of 0.94 h, which is in excellent agreement with batch data. Certainly, our batch and on-chip growth rates are comparable, and from our extensive microfluidics experience we observe with confidence that the on-chip cells are healthy and expressing normally.

***Milestone 2.5:*** *Prepare isolated RNA for Illumina sequencing using vendor provided reagents and carry out sequencing:* Completed in Q3. We have prepared libraries and carried out sequencing for all toxins of interest.

After exposing bacterial cultures to various levels of toxins, we extracted the cellular RNA and prepared it for sequencing. First, we thawed the cell pellet, homogenized the pellet using bead beating with RNase-free zirconium oxide, and extracted the RNA using Qiagen RNeasy Kits. Second, we removed any contaminating genomic DNA with DNase and performed a final purification step using a Zymo Clean and Concentrator column. Third, we prepared the RNA library for sequencing by enriching for mRNA using Epicentre Ribo-Zero rRNA removal kits. Fourth, we generated cDNA and prepared an indexed Illumina library using NEBNext Ultra Directional RNA Library Prep Kits, which retain strand-specific orientation information. Finally, we prepared and loaded these libraries on our Illumina MiSeq instrument for 2×75 bp paired-end sequencing.

| Toxin | Number DE Genes | |
|---|---|---|
| | Low Conc. | High Conc. |
| Arsenic | 3 | 14 |
| Cadmium | 132 | 38 |
| Chromium(VI) | 270 | 377 |
| Cobalt | 609 | 1274 |
| Copper | 110 | 77 |
| Iron | 32 | 70 |
| Lead | 99 | 489 |
| Malathion | 370 | 41 |
| Mercury | 165 | 218 |
| Nickel | 43 | 95 |
| Zinc | 24 | 596 |

*Table 6: Number of differentially expressed (DE) genes at a false discovery rate below 1% for each toxin exposure condition of interest.*

***Milestone 2.6:*** *Analyze data to determine candidate genes which are induced upon toxin exposure:* Completed in Q3. We analyzed our sequencing data and identified significant numbers of candidate genes for each toxin of interest at false discovery rates below 1%. As we explain in Milestone 3.3, we experimentally validated selected candidate genes for cobalt detection.

Our analysis of the sequencing data from RNA-Seq experiments to determine candidate genes that are induced upon toxin exposure consisted of three main tasks: sequence alignment, quantification of gene expression, and identification of differentially expressed (DE) genes. We used and rigorously tested established software and also developed our own software to corroborate and verify all the candidate DE genes (see Table 6). We took into account the stranding of the MiSeq reads to achieve greater than 95% alignment with the MG1655 genome. Details of our quantification of the DE genes are provided in Appendix B.

The resulting genes were cross-inspected using different methods to confirm that the most specific genes were selected. In order to simultaneously analyze mean fold change with respect to the
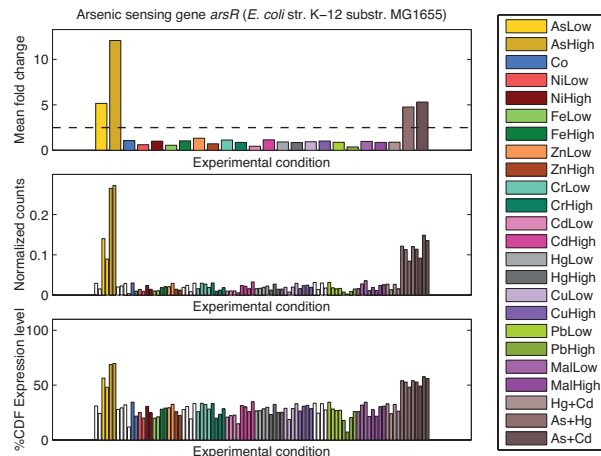


*Figure 14: Response of the arsR promoter in* **E. coli** *MG1655 for specifically sensing arsenic. Note that the arsR promoter responds specifically and monotonically to increasing concentrations of arsenic alone and in combination with other heavy metals (rightmost conditions) without exhibiting crosstalk. Plots represent the mean fold change, the normalized counts (taking into account both the library depth and the gene length), and the "%CDF expression level." This measure represents the position of the gene in the Cumulative Distribution Function (CDF) obtained from the normalized counts for a given experimental condition; that is, a large CDF indicates a high expression level for the gene compared to the other genes in the same experiment. Replicate samples for each toxin are shown in the same color. Control samples for each batch of experiments are shown in white.*

negative conditions, the normalized counts, and the expression level with respect to other genes, we implemented a toolbox that depicts a summary figure for a gene candidate as shown in Figure 14 (see the full set of gene candidates in Appendix B). From an analysis of all gene candidates, we set our mean fold change threshold to indicate DE to 2.5.

Our most sensitive and specific gene candidates for eight toxins identified via RNA-Seq analysis in *E. coli* MG1655 are shown in Figure 15. The arsR promoter in Figure 15a responds specifically and monotonically to increasing concentrations of arsenic alone and in combination with other heavy metals without exhibiting crosstalk.
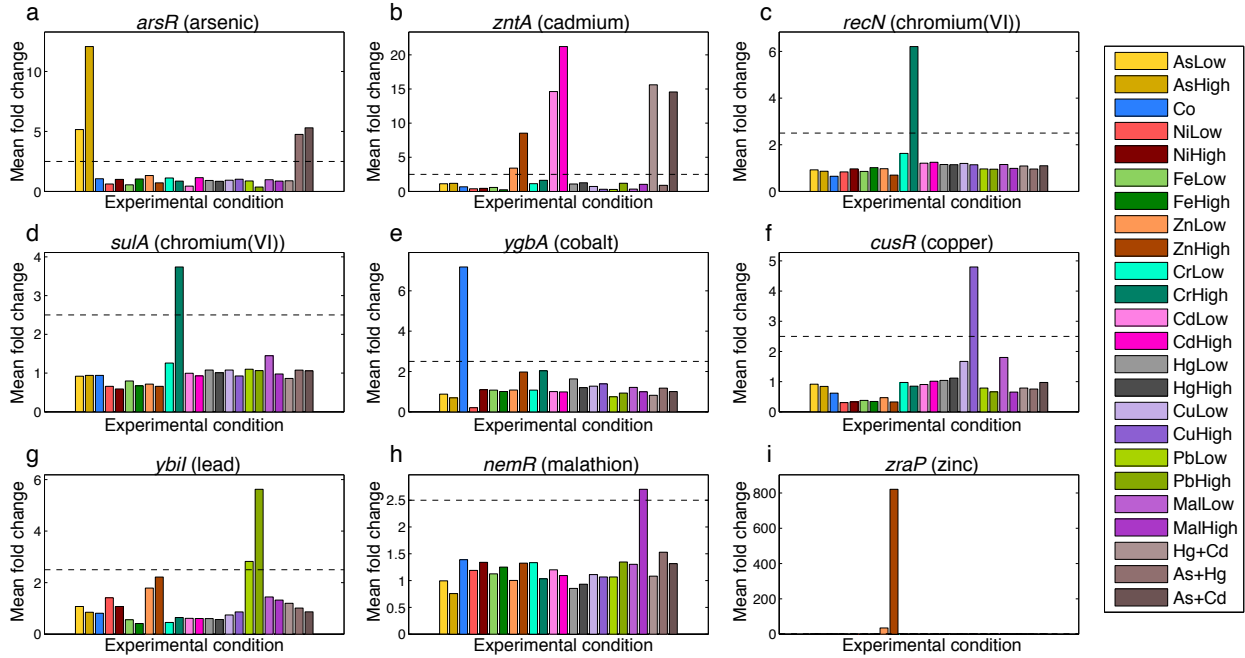
*Figure 15: Sensitive and specific gene candidates for eight toxins identified by RNA-Seq analysis in* **E. coli** *MG1655, where the mean fold change threshold to indicate DE has been set to 2.5.*

This gene serves as an excellent positive control for our analysis methods, as our literature search revealed it to be the most widely studied gene for sensing arsenic in *E. coli*. The zntA promoter in Figure 15b responds monotonically to increasing concentrations of cadmium alone and in combination with other heavy metals without exhibiting crosstalk, but it also shows sensitivity to low and high concentrations of zinc. However, since the zraP promoter is specific for zinc, the boolean logic expression (zntA)∧(¬zraP) represents a unique combined promoter response that is specific for cadmium. The recN and sulA promoters in Figure 15c–d respond specifically to high concentrations of chromium(VI) alone. These promoters are known to activate during multiple double-strand DNA breaks and the cell SOS response, respectively. Their differential expression here corroborates recent reports of chromium(VI) acting as a carcinogen in water supplies[19,20,21] and may provide information on its mode of action. The ygbA promoter in Figure 15e responds specifically to cobalt alone. The cusR promoter in Figure 15f responds specifically to high concentrations of copper alone. This gene serves as another excellent positive control, as our literature search showed it to be among the most widely studied genes for sensing copper in *E. coli*. The ybiI promoter in Figure 15g responds specifically and monotonically to increasing concentrations of lead alone and shows higher specificity than the pbrR promoter identified from the literature. The nemR promoter in Figure 15h responds specifically to high concentrations of malathion alone. The identification of a sensitive and specific gene for detecting malathion in *E. coli* is entirely unexpected, as the mechanism of toxicity for organophosphate pesticides is known to only operate in eukaryotes, as organophosphate pesticides are neurotoxins. Lastly, the zraP promoter in Figure 15i responds specifically and monotonically to

22
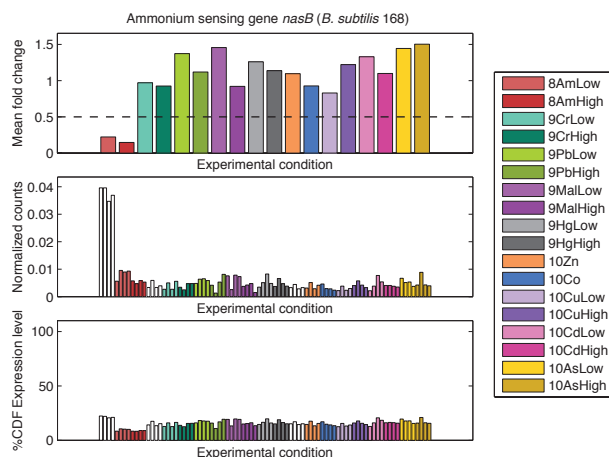
*Figure 16: The **nasB** promoter in* **B. subtilis 168** *was found to be sensitive and specifically down-regulated in response to ammonium, given a mean fold change threshold of 0.5.*

increasing concentrations of zinc alone. The sensitivity of zraP to zinc is extremely high, making this promoter an excellent candidate for logical combination with promoters exhibiting zinc interference to generate toxin-specific multi-promoter responses.

We used RNA-Seq to verify the nasB promoter in *B. subtilis* 168 as a sensitive and specific promoter for the last remaining toxin, ammonium, in early Q3 (see Fig. 16). As opposed to other response circuitry, nasB is down-regulated in response to the toxin, allowing the construction of a "lights-off" sensor. This is desirable because, unlike the other water toxins of interest, ammonium is expected to increase the growth rate of biosensor cells. In this situation, a "lights-on" sensor could be triggered by any substance that increases the growth rate and resulting GFP production ability of cells, whereas a "lights-off" sensor remains immune to this effect.

In summary, **through constructing strains based on literature searches and performing RNA-Seq analysis, we discovered sensitive promoters for all toxins of interest**. See Appendix A for a full list of candidate toxin-responsive promoters.

## Task B

**Task Objective:** Develop a preliminary microfluidic device that can culture many independent *E. coli* sensor strains simultaneously

**Milestone:** 12 months

**Metrics/Completion Criteria:**

Milestone 3: Library expansion of transcription based sensors

Milestone 4: Microfluidic device development to support environmental sensing

Milestone 5: Low cost optical methods development

**Deliverable:** Library of *E. coli* clones with target promoters producing GFP; preliminary microfluidic device capable of culturing many different clonal populations in a defined array

## Reporting Updates

**Milestone 3: Library expansion of transcription based sensors**

*Summary:* We have completed this milestone. The promoter regions for the differentially expressed genes identified in Milestone 2 were located using computational tools (Milestone 3.1). We have cloned these promoters into the Milestone 1 expression systems (Milestone 3.2). We validated the functionality of these new promoter constructs using microfluidics (Milestone 3.3). This validation of our RNA-Seq identified promoters provides strong evidence of the power and utility of our approach to developing novel transcription based biosensors.

**Milestone 3.1:** *Perform sequence analysis to determine the promoters of candidate genes identified in Milestone 2:* Completed in Q2. We have identified the promoter regions for all candidate genes. The promoters of in-
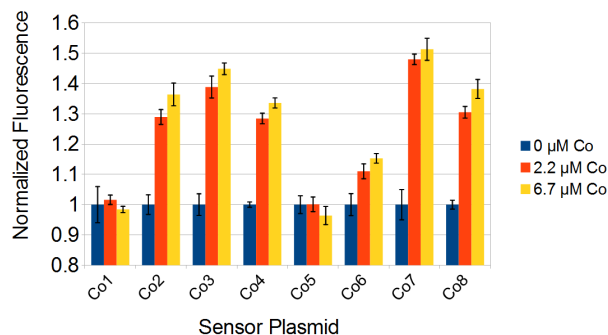
*Figure 17: Validation of five cobalt-sensing promoters (Co3, Co4, Co6, Co7, Co8) identified via RNA-Seq analysis with* E. coli *and one (Co2) identified from literature.*

terest are located on the genome directly upstream of the genes identified by RNA-Seq. For uncharacterized promoters, we used 200bp upstream of the gene's transcription start site to construct the sensor plasmid. Including this entire region ensures that even cryptic regulatory sites acting on the promoter are included in the sensor construct.

***Milestone 3.2:*** *Clone the promoters into the expression system validated in Milestone 1:* Completed in Q2. We have cloned all identified promoters into the standardized plasmid expression system validated in Milestone 1.

***Milestone 3.3:*** *Validate the response of the promoters using the methods developed in Milestone 1:* Completed in Q2, with updated results through Q4.

In Q2, we validated our RNA-Seq approach to identifying toxin-responsive promoters using plate reader data from five cobalt-sensing promoters identified via RNA-Seq analysis (see Fig. 17). We grew strains in microplate wells in the presence of various concentrations of cobalt and measured GFP fluorescence. *E. coli* MG1655 strains Co1 and Co2 incorporate the nmtR promoter from the *M. tuberculosis* genome as the cobalt-sensing ele-

ment, whereas *E. coli* MG1655 strains Co3–Co8 incorporate cobalt-responsive promoters identified through RNA-Seq. Odd-numbered strains express GFP using the native RBS, whereas even-numbered strains use the optimized Lutz RBS. We concluded that strains Co2–Co4 and Co6–Co8 significantly respond to the presence of cobalt in the growth medium. Therefore, **we have demonstrated that our exhaustive methodology for RNA-Seq analysis supported by multiple techniques and algorithms works**.

The set of validated constructs shown in Figure 18 contains one construct that has been demonstrated sensitive to each toxin of interest within our sensor prototype, with the exception of malathion for which the positive results that we observed in our original microscopy runs (shown in this plot) have not yet translated to our sensor boxes. During investigation, we discovered a more robust malathion detection scheme that leverages its esterase inhibition effect in yeast (see Fig. 6). Our successful testing of this method on our microscope illustrates the power of our forthcoming whole-genome detection scheme, whereby we will continually measure the responses of 2,000 *E. coli* and *S. cerevisiae* promoters arrayed in a microfluidic chip. We expect the observed esterase inhibition to be significantly represented in the genomic response of this multi-strain library during malathion exposure.

**Milestone 4: Microfluidic device development to support environmental sensing**

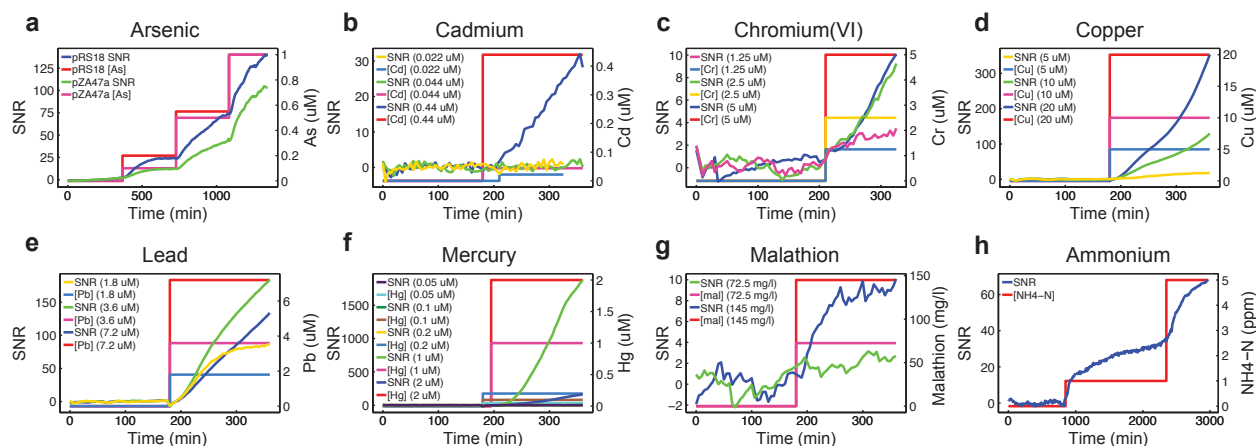*Summary:* We have successfully completed this milestone. We developed a concentrated media additive to enable robust growth of *E. coli* in natural water sources (Milestone 4.1). We cultured *E. coli* in our microfluidic devices for up to 50 days, demonstrating long-term reliability (Milestone 4.2). We developed and op-

*Figure 18: On-chip time-lapse induction responses of promoter constructs on our research grade microscope. All sensor constructs are on plasmids transformed into* **E. coli** *MG1655, except for the ammonium construct which is integrated into the genome of* **B. subtilis** *NCIB 3610. a) pRS18 and pZA47a show induction by arsenic. b) Cd1 shows induction by cadmium. c) Cr11 shows induction by chromium(VI). d) Cu1 shows induction by copper. e) Pb7 shows induction by lead. f) Hg3 shows induction by mercury. g) Mal1 initially showed induction by malathion; however, we have not been able to acquire similar data in our sensor boxes. We have recently identified a more robust sensing mechanism using the esterase inhibition effect of malathion in yeast. h) Amm3 shows induction by ammonium.*

timized a microfluidics-capable peristaltic pump for extracting natural water (Milestone 4.3). We greatly expanded our microfluidic device, allowing the culture of multiple strains (Milestone 4.4). We significantly increased the fluorescence signal from our microfluidically cultured cells by optimizing the trap geometry (Milestone 4.5). We developed techniques to prevent the release of our genetically modified organisms (Milestone 4.6). We developed a reliable method to freeze-dry and store cells at room temperature (Milestone 4.7). We modified our "gill" microfluidic device to allow freeze-drying of cells on-chip (Milestone 4.8). We successfully tested our microfluidic devices using our low-cost optical system (Milestone 4.9).

***Milestone 4.1:*** *Develop a concentrated media additive that can be mixed with natural waters to support* E. coli *growth:* Completed in Q2, with updated results through Q3. We have developed and tested medium formulations that successfully culture *E. coli* and *B. subtilis* in batch and on-chip.

For all batch culture experiments, M9 minimal medium supplemented with 0.4% w/v glucose, 0.1 mM $CaCl_2$, and 2 mM $MgSO_4$ was used. For *B. subtilis*, the medium was additionally supplemented with 0.075% v/v TWEEN 20, 50 $\mu$M $FeCl_3$, 50 $\mu$M $MnCl_2$, and 1 $\mu$M $ZnCl_2$. For *B. subtilis* ammonium exposure experiments, the $NH_4Cl$ in M9 minimal medium was replaced with $NaNO_3$, keeping the concentration of nitrogen constant.

For microfluidic experiments, we developed a minimal medium optimal for growth of bacteria and heavy metal sensing, adapted from HMM. This medium replaces the inorganic phosphate in M9 minimal medium with glycerol-2-phosphate, MOPS (pH = 7.2), and KCl. Inorganic phosphate is undesirable because of its metal chelation properties and its propensity to form calcium phosphate deposits within microfluidic channels. To minimize contaminating metals, all microfluidic experiments were carried out with media made with extra high purity salts where available. We found that when using these pure

salts robust *E. coli* growth required supplementing the medium with iron, and robust *B. subtilis* growth required iron, zinc, and manganese.

The final composition of our *E. coli* medium, following on-chip mixing with source water, was:

1. 40 mM PharmaGrade MOPS [Sigma #PHG0007-1KG] (from 1 M stock at pH 7.2)
2. 4 mM glycerol-2-phosphate [Sigma #G6501-25G]
3. 0.4% w/v dextrose (glucose) [Sigma #D9434-1KG]
4. 1 g/l (19 mM or 262 ppm $NH_4$-N) TraceSelect $NH_4Cl$ [Sigma #09725-100G]
5. 3.7 g/l (50 mM) TraceSelect KCl [Sigma #05257-100G]
6. 0.075% v/v TWEEN 20 [Acros Organics #23336-0010]
7. 50 $\mu$g/mL spectinomycin (from spectinomycin dihydrochloride pentahydrate) [Sigma #S4014-5G]
8. 1 $\mu$M $FeCl_3$ [Alfa Aesar #A16231-500G]
9. 0.01 mM $CaCl_2$ [Macron Fine Chemicals #4160-12]
10. 0.2 mM $MgSO_4$ [Macron Fine Chemicals #6066-04]

The final composition of our *B. subtilis* medium, following on-chip mixing with source water, was the same as for *E. coli* with the following modifications:

1. Replace $NH_4Cl$ with 1.6 g/l $NaNO_3$
2. Use 50 $\mu$M $FeCl_3$ instead of 1 $\mu$M
3. Add 50 $\mu$M $MnCl_2$ [Baker #2540-01]
4. Add 1 $\mu$M $ZnCl_2$ [Macron #8780-04]

Note that TraceSelect formulations of reagents were used when available to minimize the potential for heavy metal contamination of the media.

***Milestone 4.2:*** *Perform testing to demonstrate the long term culture stability of such cells. Quantify growth rate of cells and lifespan of cultures:* Completed in Q2, with updated results through Q4. We have successfully grown cells in microfluidic devices with a stable growth rate comparable to batch culture.

We have continuously grown *E. coli* MG1655 in the gill chip for 23 days using growth medium concentrate mixed with natural water from Lake Miramar. In the most recent experiment, 10× M9 medium concentrate was mixed with Milli-Q water in a 1:9 ratio at a total flow rate of 1 ml/h using a dual-channel Instech P625 peristaltic pump. Cells were observed to grow and express GFP in the traps after 50 days in the device, proving that the sensor strain is stable over this time period.

***Milestone 4.3:*** *Develop a metering and mixing system to combine the concentrated media stock with natural water at a fixed ratio and mix it well before cell exposure:* Completed in Q2, with updated results through Q4. We used an Instech P625 peristaltic pump to mix concentrated media with natural water at a fixed ratio and calibrated the pump to achieve the desired flow rate. Since our media additive stocks are concentrated 10–20×, we dilute them with natural water before delivery to the cells by using a peristaltic pump to drive each liquid at a defined volumetric flow rate through silicone tubing into the chip. Our pumping scheme uses a custom-made dual-channel tubing set with different tubing inner diameters, which results in a constant flow ratio. In laboratory tests, we have used tubing sets with 1:20 and 1:9 flow ratios to successfully combine and completely mix the two liquid streams on-chip using staggered herringbone mixers to support healthy cell growth.

***Milestone 4.4:*** *Expand the current* E. coli *large scale culture device (biopixel device) to have 500 individually addressable chambers:* We have designed a large-scale gill chip using nested symmetrically-split parallel channels to produce balanced flow (see Fig. 19). However, our analysis of RNA-Seq data determined that the incidence of specific toxin-sensing promoters in *E. coli* is
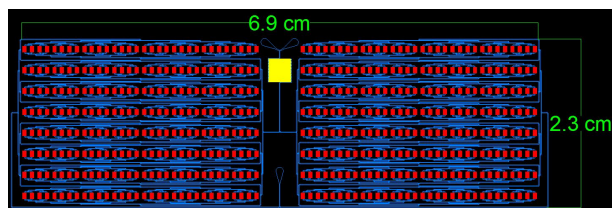
*Figure 19: CAD design for a highly parallel microfluidic device capable of culturing 512 unique toxin-sensing promoter strains in individually-addressable "gill" cell-trapping regions. Two inlet ports at the top are combined and mixed during passage through staggered herringbone mixers (yellow) before branching into 512 "gill" trapping regions (red) and recombining at the outlet port. The footprint of the entire device is small enough to fit on a 1"×3" glass slide.*

greater than expected. Therefore, we do not require hundreds of cell-trapping regions in the microfluidic device, and we decided to pursue a cell patterning strategy whereby cells are micropipetted into shallow but relatively large reservoirs upstream of the cell trapping regions.

***Milestone 4.5:*** *Increase the height of the Biopixel device's chambers to increase the output optical signal. Ensure the cell growth dynamics are unchanged:* Completed in Q1. We constructed and tested an optimized gill chip design that raises the height of the cell trapping regions to 50 $\mu$m, thereby increasing the fluorescent signal.

Notably, we have observed synthetic strains of *E. coli* MG1655 produce levels of GFP within the device that are visible to the naked eye using the appropriate filters. Over the course of dozens of experiments, we have verified that cell growth rate is unaffected by these minor modifications to the "biopixel" cell trap design.

***Milestone 4.6:*** *Develop a UV LED system that kills the cells as they emerge from the trapping region and enter waste collection. Perform viability test on media exiting the waste trap:* We concluded that UV LEDs are an inefficient method for sanitizing the chip effluent.

Short-wave UV LEDs consume power, emit heat, and have short lifetimes, which unnecessarily reduces the operational lifetime of our biosensor device. Instead, we decided to route cell waste to a reservoir containing bleach, which has been used as an effective sanitizer for hundreds of years.

In Q2, we tested this strategy by depositing effluent from the pump mixing experiments into a 1-l glass bottle pre-filled with 100 ml of bleach. The waste bottle was used to collect cell waste throughout four experiments over a period of two months and filled to around 500 ml. To test for cell viability, the waste bottle was stirred, and 20 $\mu$l of the contents was plated on LB agar. As a positive control, 45 $\mu$l of waste was removed, bleached again, and plated using the same method. Neither the pre-bleached waste nor the post-bleached waste showed any bacterial growth after 1 day on LB agar without antibiotics, suggesting the pre-bleaching method is sufficient to eliminate viable cells in the waste bottle.

***Milestone 4.7:*** *Develop a method to freeze-dry cells allowing them to be rehydrated with little loss of viability:* Completed in Q3. We have successfully developed a method for in-chip lyophilization and revival after long-term storage.

Several cryoprotectants suitable for engineered biosensor strains were formulated from a combination of literature-based protocols, current industrial practices, and experimentation. The investigated cryoprotectants include:

1. 2.5% Luria-Bertrani Broth (LB) (w/v) + spectinomycin
2. 2.5% LB + 0.4% glucose (w/v) + spectinomycin
3. 2.5% LB + 0.4% sucrose (w/v) + spectinomycin
4. 2.5% LB + 0.4% trehalose (w/v) + spectinomycin
5. M9 + 0.4% glucose + spectinomycin

Relative cryoprotectant efficacy was de-
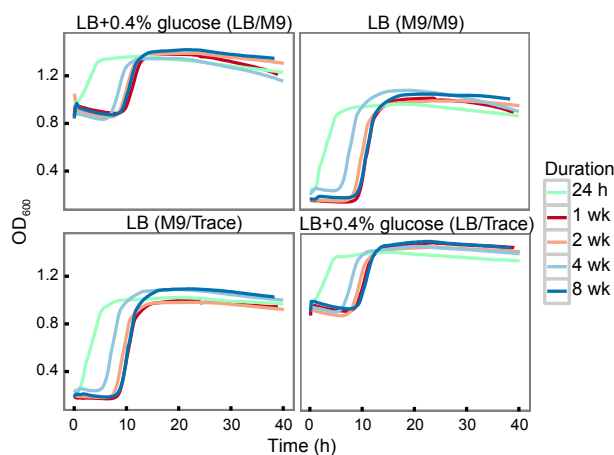
*Figure 20: Cryoprotectant revival rates for arsenic-detecting* E. coli *plasmid/strain pLB-As3/MG1655. The cryoprotectant is listed over each plot, with the overnight and revival medium listed in that order in parentheses (*i.e. *a strain cryoprotected in LB after being grown overnight in M9 and revived in Trace Select M9 medium is denoted by "LB (M9/Trace)").*



*Figure 21: Images showing the on-chip revival of lyophilized sensor strain pLB-As7 (MG1655). Device image a) 0.5 h after chip wetting and b) after 11 hours of robust post-revival growth. The strain was loaded, grown to confluence, lyophilized in LB + 0.4% sucrose + spectinomycin for 17 h (9 h freeze, 8 h dry), stored for 24 h, and revived in LB + spectinomycin.*

termined via plate reader revival experiments performed 24 h, 1 wk, 2 wk, 4 wk, and 8 wk after lyophilization. Cells were revived via rehydration and resuspension in 200 $\mu$l of revival medium within microplate wells. The plates were then immediately placed into a Tecan Infinite M200 Pro plate reader, where growth rates were monitored over the next 48 h.
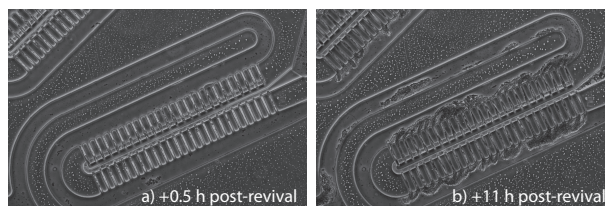
Revival media included:

1. M9 + 0.4% glucose + spectinomycin
2. Trace Select M9 + 0.4% glucose + spectinomycin
3. Trace Select M9 + 0.4% glucose
4. HM9 (nitrate) + 0.4% glucose

and were selected to be representative of the growth medium used in the final device.

Strains protected with optimal cryoprotectants showed little difference in viability between cryoprotectants after two months of preservation. Both *E. coli* and *B. subtilis* strains responded similarly to lyophilization in the cryoprotectants listed above (see Fig. 20).

The best cryoprotectant was found to be LB + 0.4% sucrose and was used to test batch and on-chip lyophilization and revival in biosensor host strains *E. coli* MG1655 and LABEC31 and *B. subtilis* NCIB 3610. In batch, revival efficacy experiments were performed at 24 h, 1 wk, 2 wk, 4 wk, and 8 wk after lyophilization, with no observed reduction in viability. For on-chip testing, strains were cultured overnight to stationary and sporulation phases, respectively. Strains were then double-washed in cryoprotectant and concentrated to 50$\times$ their batch culture concentration before being injected into an 18-strain chip through independent loading ports. Loading ports were sealed with a fast-curing silicone elastomer (Sylgard 170, Dow Corning), then the device was lyophilized in a commercial freeze-dryer for 17 h before being nitrogen-flushed, desiccated, and sealed within opaque packaging.

Following room-temperature storage for up to two months, strains were revived via de-gas driven chip wetting and pump-driven flow. Upon introducing fresh medium, strains revived on time scales equivalent to those in batch (see Fig. 21 and movie here). In addition, the ammonium-sensing *B. subtilis* NCIB 3610 strain was successfully sporulated using standard sporulation medium, which offers an additional

preservation method with extreme shelf-life.

***Milestone 4.8:*** *Develop a deposition technique to place cells into a region of a microfluidic device that is then bonded to a glass coverslip:* Completed in Q3. We have successfully developed a deposition strategy whereby individual biosensor strains are injected into on-chip reservoirs, where they are lyophilized, stored, and revived.

To independently culture multiple biosensor strains, an 18-strain chip was designed, built, and successfully tested for multi-strain loading and freeze-drying. The chip dimensions conform to the requirements of both the strains and the biosensor's optical detection system.

***Milestone 4.9:*** *Develop and test the microfluidic device in a laboratory environment:* Completed in Q2. We have conducted extensive testing of our microfluidic devices in a laboratory environment, using both research grade and low cost optical systems.

In Q2, the microfluidic "gill" chip described in Milestone 1.4 was used to test the optical system developed by the Ziva Corporation. We compared the optical system to that of our research grade microscope, an Olympus IX81. The Ziva optical system was designed with lower resolution optics compared to those of the $4\times$ objective on the IX81 in order to lower cost while increasing the imaged field of view. Although the Ziva optics are lower resolution, the produced images compare favorably with the IX81 as shown in Figure 22.

**Milestone 5: Low cost optical methods development**

*Summary:* We have successfully completed this milestone. We have developed a low cost optical system by partnering with the Ziva Corporation, thereby completing sub-milestones 5.1, 5.2, and 5.3. We have also determined that bioluminescent
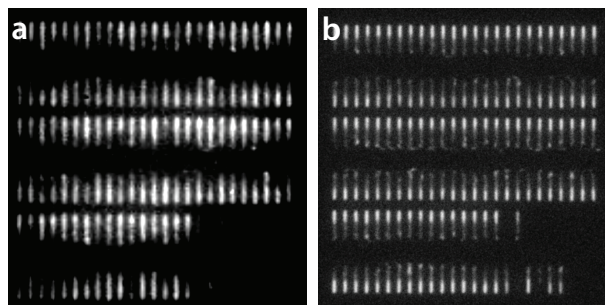


*Figure 22: A comparison of optical systems. The same microfluidic device was imaged with a) low cost biosensor optics developed by the Ziva Corporation and b) a research grade Olympus IX81 microscope. While the Ziva image may appear inferior, this is an artifact due to zooming of the image for comparison (the Ziva system is designed with a wide field of view compared to the Olympus). We have verified that the Ziva system yields images that equal the quality of the Olympus system, which is remarkable given the difference in cost ($2K for Ziva versus $125K for Olympus).*

systems can be significantly more sensitive than fluorescence based systems, thus completing Milestone 5.4.

***Milestone 5.1:*** *Design and construct an LED array for the GFP excitation of the cells in the microfluidic device:* Completed in Q2. The excitation LED system has been designed and delivered.

In Q2, the optical setup was designed in partnership with the Ziva Corporation. We chose a CREE XTE Indus Star 1 Royal Blue High Power LED (manufacturer part #CREEXTE-ROY-1) with an emission wavelength maximum of 465 nm, which is well-suited for GFP excitation. The LED is paired with a 1400 mA BuckBlock constant current LED driver (manufacturer part #0A009-D-V-1400) and a LED light housing with a 15 W heat sink (manufacturer part #ALK-LH-15W).

***Milestone 5.2:*** *Design and construct optical filters for GFP excitation and emission which cover the entire area of the microfluidic device:* Completed in Q2. The optical filters

for GFP excitation and emission have been designed and delivered.

The optical setup was designed in partnership with the Ziva Corporation. The necessary filters were purchased from Semrock, Inc. The part numbers are FF495-Di03-25x36, FF01-520/35-23.3, and FF01-457/50-25 for the dichroic, emission, and excitation filters, respectively.

***Milestone 5.3:*** *Obtain and characterize the performance of a low cost camera system to image the fluorescence signal of the device:* Completed in Q3. We have built an imaging system using the "Chameleon" camera (part #CMLN-12S2M-CS) from Point Grey Research, Inc. with the design assistance of the Ziva Corporation. This is a 1.3 megapixel monochrome camera featuring a Sony ICX445 CCD imager. It contains a 12-bit analog-to-digital converter with a maximum gain of 24 dB. The camera package includes a software development kit (SDK), known as FlyCapture, which is compatible with the PandaBoard single board computer system that we have chosen for our electronics platform.

The microfluidic "gill" chip developed in Milestone 1.4 was used to compare the optical system developed by the Ziva Corporation with our research grade microscope, an Olympus IX81. The Ziva optical system was designed with lower resolution optics compared to those of the $4\times$ objective on the Olympus in order to lower cost while increasing the image field of view by $20\times$ for imaging multiple "gill" trapping regions. Images acquired with the Ziva optics compare favorably with those acquired with the Olympus, as shown in Figure 22. Notably, the Ziva optics are $\sim 2\times$ more sensitive at detecting GFP than the Olympus optics (SNRs are 23.8 and 11.6, respectively). Because our primary objective is to detect weak signals, the Ziva system outperforms the Olympus system at $\sim 50\times$ lower cost.
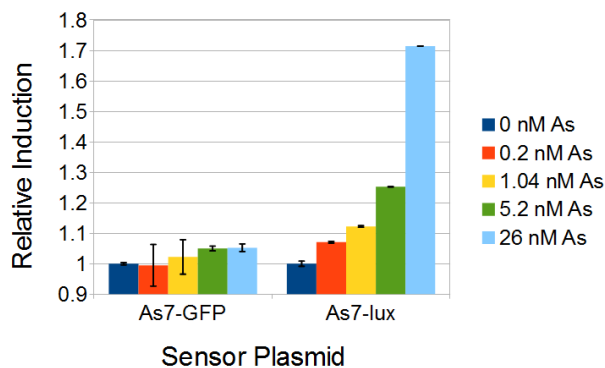


**Figure 23:** *A side-by-side comparison of reporter systems on otherwise identical plasmid backbones demonstrates the superior detection limits achieved by using luminescence.*

***Milestone 5.4:*** *Replace the GFP fluorescence system with a luminescent system based on the* lux *operon of* A. fischeri: Completed in Q2. We have replaced the GFP fluorescence system with a Lux system for our best-performing arsenic sensor plasmid.

The arsenic sensor construct As7 was modified to replace the *gfp* gene with the *luxCDABE* operon, and relative induction (fluorescence or luminescence) of the two constructs was tested side-by-side using a Tecan Infinite M200 Pro plate reader (see Fig. 23). The background noise was significantly lower for the Lux construct, allowing the detection of arsenic at a concentration of 0.2 nM. In contrast, the GFP construct was sensitive to arsenic only above a concentration of 26 nM. For comparison, we were able to detect an arsenic concentration of 130 nM using the GFP construct in a microfluidic device. Thus, replacing GFP with Lux promises to greatly increase the sensitivity of the biosensor.

## Task C

**Task Objective:** Develop a prototype of a deployable device for cheap and continuous monitoring of water contamination by specific target compounds

**Milestone:** 12 months

**Metrics/Completion Criteria:**

Milestone 6: Device prototype development

Milestone 7: Build computational models to determine the threshold of detection for specific sensors based on models of experimental GFP responses

Milestone 8: Use the models obtained in Milestone 7 to investigate whether the combination of nonspecific sensor responses to some toxins can be utilized to improve specificity

Milestone 9: Develop the controller board to carry out sophisticated analysis of complex data

**Deliverable:** Finalized device prototype and computational algorithm for continuous water quality monitoring

## Reporting Updates

**Milestone 6: Device prototype development**

*Summary:* We have successfully completed this milestone. We obtained and characterized a low power peristaltic pump (Milestone 6.1). We have obtained a filter to prevent device clogging (Milestone 6.2). We have demonstrated that flow reversal to prevent clogging is not necessary based on our filtration design (Milestone 6.3). We have obtained an environmental enclosure and heater (Milestones 6.4 and 6.5). We have obtained and developed the software components necessary for device control and data transmission (Milestone 6.6). We have developed a positive control stock for sensor testing purposes (Milestone 6.7). We have developed a solar cell based charging system (Milestone 6.8). We have assembled a fully functional prototype and have tested it in an outdoor environment (Milestone 6.9).

*Milestone 6.1:* *Find and characterize a low power pumping system that can pump natural water through the microfluidic device at a rate of 1 ml/h:* Completed in Q1. We have purchased and tested a low power peristaltic pump compatible with microfluidics.

*Milestone 6.2:* *Develop a filter system to prevent clogging of the microfluidic device:* Completed in Q1. We have successfully implemented filtering to prevent clogging when pumping natural water sources.

*Milestone 6.3:* *Develop a flow reversal pumping regime to help clear the filter of contaminants. Test and revise the de-clogging method to ensure a runtime of at least one month per sensor:* Completed in Q1. Due to the success of our filter system developed in Milestone 6.2, we determined that flow reversal was not necessary.

*Milestone 6.4:* *Develop a waterproof enclosure that can be used to house the device and the electronics:* Completed in Q1. We have purchased and tested an all-weather enclosure suitable for outdoor use.

*Milestone 6.5:* *Develop a heating system to ensure the microfluidic device maintains the appropriate growth temperature:* Completed in Q1. We have tested a heating system for our device.

*Milestone 6.6:* *Develop the electronics/software to coordinate the pumping regimes, image capture, data transmission and device power management:* Completed in Q2. We have obtained the necessary hardware and software for controlling the electronic components of our biosensor.

We assembled components to support the time-lapse fluorescence imaging and analysis of cells within microfluidic devices. We also implemented a low cost, low power, all-solid-state PandaBoard single-board computer to control image acquisition and peripherals. We implemented a

fluorescence imaging assembly designed by the Ziva Corporation (Fig. 25d) to mount in our prototype imaging scaffold. Light from a blue excitation LED reflects off a dichroic mirror and illuminates the sample. The optics collect fluorescent light emitted by the sample and pass it though the dichroic mirror and emission filter while focusing it to an image at the CCD camera. We developed a custom software package to control data acquisition, regulate pump speed, and maintain appropriate environmental conditions for the biosensor.

***Milestone 6.7:*** *Develop a chemical control stock to test the proper operation of the biosensor (*i.e. *add low doses of toxins for positive control):* Completed in Q1. We developed concentrated chemical stocks of the eight toxins that are easily mixed to form a solution containing low doses of toxins to serve as a positive control in testing proper biosensor operation.

***Milestone 6.8:*** *Develop a solar powered version of the prototype, including battery panels and charge controller):* Completed in Q2. We have developed an independent solar charging station that can easily interface with our biosensor device.

To ensure compatibility with solar or battery power, we selected all components of our device (including the heater) to be DC powered. We then located and purchased an isolated DC/DC converter (Mean Well part #SD-50A-12) that is capable of generating a regulated 12 VDC output from an unregulated 9.2–18 VDC input. The regulated 12 VDC supply is used to power our pump, heater, and electronics. The unregulated 9.2–18 VDC input is within the voltage range of a standard lead acid battery; therefore, we purchased a deep cycle, lead acid battery (Interstate battery part #DCM0035) with sufficient energy storage (35 Ah) to power our device for approximately two days. To charge the battery, we purchased a 100 W monocrystalline solar panel and a 30 A pulse wave modulation (PWM) charge controller from Renology Solar. The solar panel components were assembled and mounted onto a galvanized steel pole at our field test facility.

***Milestone 6.9:*** *Develop a completed device prototype and test in an outdoor environment):* Completed in Q4. We have assembled the individual components listed in sub-milestones 6.1-6.8 into a functional prototype that is capable of acquiring and processing data. Images of this prototype are shown in Figure 25.

We have finalized the design of our 18-strain microfluidic device that can collect toxin response data from 18 different toxin-specific strains simultaneously (Fig. 24). This is housed inside our sensor prototype and imaged using our custom optics. The prototype is contained in a $16'' \times 14''$ fiberglass enclosure (Fig. 25a) with an aluminum front panel designed to protect the interior components from the environment. A proportional-integral-derivative (PID) temperature controller is visible in the upper left hand corner of the front panel with the peristaltic water pump in the lower right corner. The interior of the prototype is shown in Fig. 25b with the electronics sub-enclosure (Fig. 25b-1) and Ziva optical system clearly visible (Fig. 25b-5). The use of a PID controller rather than a simple thermostat is necessary for precise control of the enclosure's internal temperature to within $0.1°C$ of the set point. Individual components are numbered 1–8. Briefly, the electronics enclosure (Fig. 25b-1) contains the hardware for controlling the Ziva optics, processing images, and transmitting data. This system can also communicate with the temperature controller (Fig. 25b-2) to adjust the interior temperature of the enclosure (normally kept at $37°C$) using the Modbus protocol.
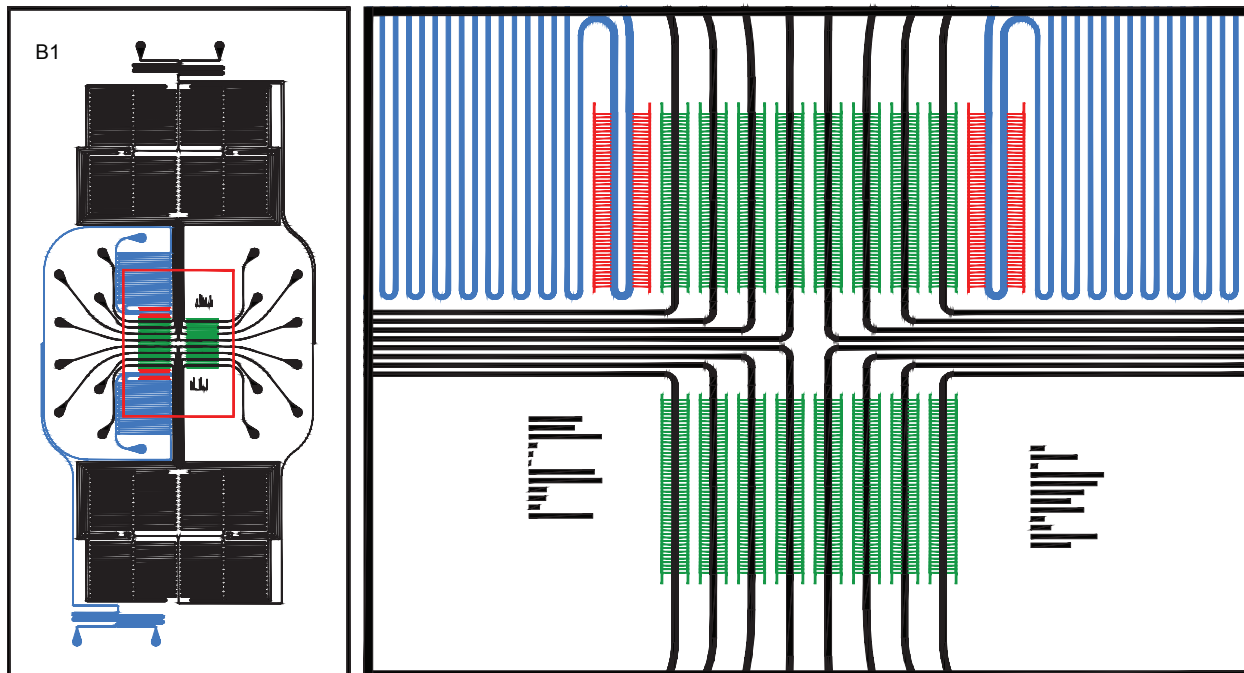
*Figure 24: CAD drawing of the 18-strain microfluidic chip used to collect toxin response data from 18 different toxin-specific strains simultaneously. Left) View of the entire microfluidic device. The resistance to each cell trapping region is identical, ensuring equivalent flow to all areas of the device. Black channels:* **E. coli** *analyte distribution channels. Blue channels:* **B. subtilis** *analyte distribution channels. Red rectangle: magnified area shown in the right panel. Right) The field of view captured by our imaging system. Red channels:* **B. subtilis** *"gill" trapping regions. Green channels:* **E. coli** *"gill" trapping regions. The collections of rectangles flanking the lower* **E. coli** *trapping region are present to assist in automatic image registration.*

This version of the prototype is designed to be powered from a 120 VAC source, and the AC power distribution block, along with a supplemental protection circuit breaker and solid state relay for controlling the heater, is shown in Fig. 25b-3. A tri-voltage power supply, outputting 5, 12, and 24 VDC for the electronics and pumps, is shown in Fig. 25b-4. The Ziva optical system for acquiring image data is shown in Fig. 25b-5. The heater is combined with a circulating fan to distribute warm air throughout the enclosure (Fig. 25b-6). Briefly, the user inputs the desired temperature through a custom software package designed by our group that runs on the PandaBoard system (Fig. 25b-1). The PandaBoard communicates with the PID controller (Fig. 25b-2) to set the desired temperature using the Modbus protocol. The PID controller modulates the heater's activity based on the set point temperature and the current temperature of the enclosure. To regulate the heater's output, the PID controller generates a pulse wave signal that drives the activity of a solid state relay (Fig. 25b-3), which turns the heater's AC power source on and off (Note: the fan remains constantly on to circulate air). Figure 25b-7 shows the DC power distribution system for the electronics and peristaltic water pump (Fig. 25b-8).

To mount the necessary electronics for acquiring and processing data and to protect them from water exposure, we designed a custom sub-enclosure using SolidWorks (Dassault Systems) and contracted its fabrication using additive manufactur-
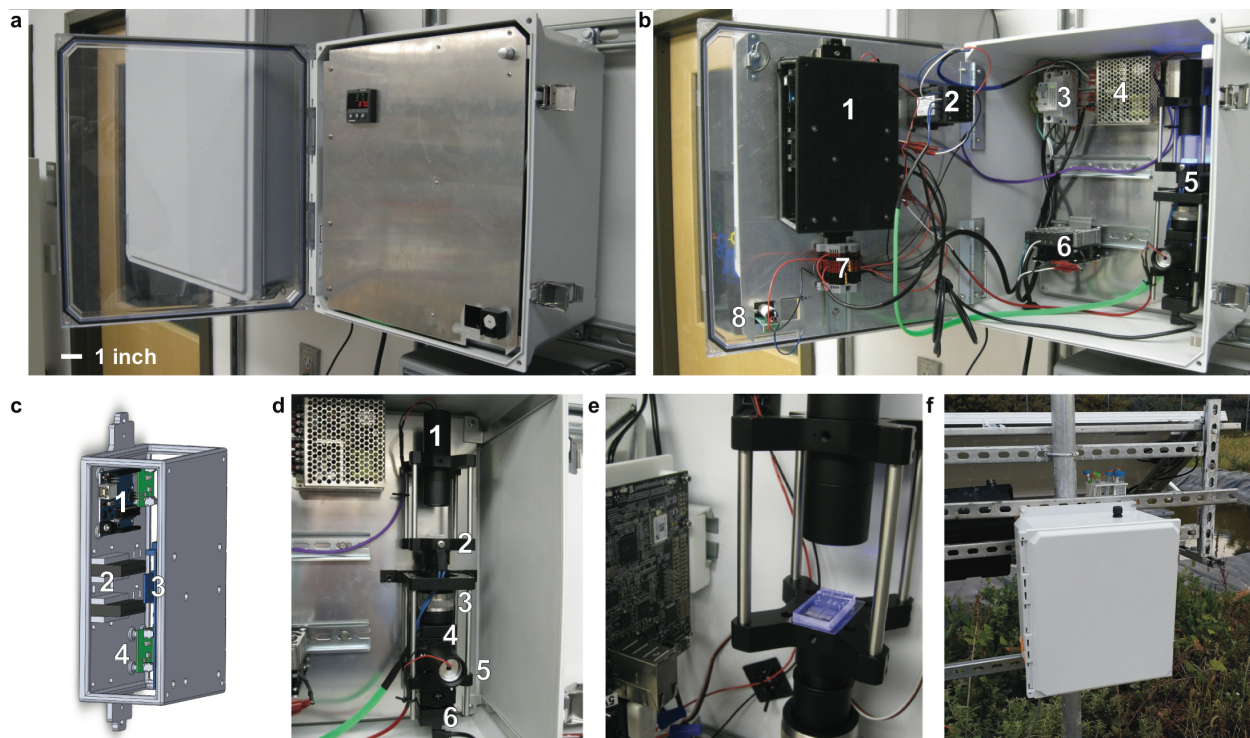
*Figure 25: Biosensor self-contained prototype: a) Device enclosure, note temperature controller (upper left) and peristaltic pump (lower right) mounted to the aluminum front panel. b) Enclosure with front panel opened to expose internal components. b1: Electronics sub-enclosure, b2: temperature controller, b3: AC power distribution devices, b4: tri-output DC power supply, b5: Ziva optical assembly, b6: fan/heater, b7: DC power distribution block, b8: peristaltic pump. c) SolidWorks detail of the electronics sub-enclosure. c1: Arduino Uno, c2: BuckBlock LED drivers, c3: PandaBoard, c4: LED control relays. d) Ziva optical assembly. d1: transmitted light optics, d2: stage/temperature probe, d3: focal adjustment knob, d4: dichroic mirror holder, d5: GFP excitation optics, d6: monochrome camera. e) Microfluidic device being illuminated with the GFP excitation LED. f) Image of the prototype running in an outdoor environment with solar power.*

ing (3-D printing) by a local machine shop (Fig. 25c). A depiction of the SolidWorks representation of this enclosure, showing an Arduino Uno (Fig. 25c-1), BuckBlock LED drivers (Fig. 25c-2), PandaBoard system on a chip (Fig. 25c-3) and LED control relays (Fig. 25c-4) is shown in this Figure. The PandaBoard communicates with the Arduino over a RS-232 (serial) link to modulate the LED control relays. The Arduino then generates a pulse wave modulated output signal that is interpreted by the BuckBlock LED drivers to control the brightness of the LEDs (one each for transmitted light and GFP excitation). The PandaBoard is then responsible for acquir-

ing an image from the Ziva optical system (Fig. 25b-5) and finally signalling for the LEDs to be turned off. The acquired data is analyzed on the PandaBoard, and the results are transmitted via a secure Wi-Fi link to our data repository server. The software to accomplish this was custom programmed in Java and C using the Point Grey FlyCapture SDK and implemented on a PandaBoard ES rev B.3 running Ubuntu Server. The various components of the Ziva optical system are shown in Figure 25d, including the transmitted light optics (Fig. 25d-1), the microfluidic device stage and thermistor temperature probe (Fig. 25d-2), the focus ad-

justment system (Fig. 25d-3), the dichroic mirror holder (Fig. 25d-4), the GFP excitation system (Fig. 25d-5), and the Point Grey Chameleon monochrome camera (Fig. 25d-6). An image of a microfluidic device, illuminated with the Ziva GFP excitation optics is shown in Figure 25e. We built five replicates of this completed prototype to run in our laboratory to collect data for the classifier. As proof of principle, we have also demonstrated that the biosensor can run in an outdoor environment and draw a sample of raw water from a public water source (Fig. 25f).

## Milestone 7: Build computational models to determine the threshold of detection for specific sensors based on models of experimental GFP responses

*Summary:* We have successfully completed this milestone. We used machine learning techniques to determine the relationships between the GFP output signal and the presence of a toxin (Milestone 7.1). We have created a database containing the collected sensor response data and have incorporated all of the collected data (Milestone 7.2). We have quantified the GFP threshold of detection for each sensor construct (Milestone 7.3). We have constructed Receiver Operating Characteristic (ROC) curves for each sensor to achieve robust sensing (Milestone 7.4).

***Milestone 7.1:*** *Characterization of the family of GFP sensor responses to the set of chemicals of interest via computational models):* Completed in Q3. We constructed machine learning models capable of inferring the relationships between the GFP sensor responses and the presence or absence of a toxin at a given concentration. The algorithm *learns* these relationships from a set of training samples (GFP sensor responses) defined by the set of experimental conditions from which they were generated. The

| Toxin | Conc. ($\mu$M) | Accuracy (%) |
|---|---|---|
| Arsenic (pRS18) | 0.2 | 97.35 |
| | 0.55 | 100 |
| | 1 | 100 |
| Arsenic (pZA47a) | 0.1 | 97.35 |
| | 0.5 | 100 |
| | 1 | 100 |
| Cadmium | 0.022 | 65.00 |
| | 0.044 | 54.16 |
| | 0.44 | 90.00 |
| Chromium(VI) | 1.25 | 96.67 |
| | 2.5 | 97.50 |
| | 5 | 94.17 |
| Copper | 5 | 95.00 |
| | 10 | 96.67 |
| | 20 | 95.00 |
| Lead | 1.8 | 95.83 |
| | 3.6 | 95.00 |
| | 7.2 | 95.00 |
| Mercury | 0.2 | 63.33 |
| | 1 | 87.50 |
| | 2 | 99.17 |
| Ammonium | 1 ppm | 71.05 |
| | 5 ppm | 91.79 |

*Table 7: Classification accuracy results obtained from the GFP sensor responses for multiple toxins at different concentrations.*

aim of the algorithm is to provide a general method capable of determining the experimental conditions associated with GFP sensor responses through the use of historical data.

Specifically, we have built classification models based on Support Vector Machines (SVMs), which is one of the most popular classifiers due to its excellent performance in many contexts and its solid mathematical basis.[22] For each toxin and concentration, we solved a binary classification problem in which the positive class represents the presence of the toxin in water and the negative class is associated with clean environments. Patterns were constructed with features containing GFP sensor responses at various timestamps to capture the tempo-

ral dynamics of the GFP signal. The optimal meta-parameters of the SVM classifier were determined by applying a 5-cross-validation during the training phase.[23]

In order to have a reliable estimate of the performance of the model when deployed in real environments, we measured its performance over a set of samples (test patterns) not seen during the training phase. The SVM's performance was determined by the percentage of test samples correctly labeled as toxin/no toxin (classification accuracy). 80% of samples were used for training the SVM models and the remaining 20% of samples was used to evaluate their effectiveness. We generated 20 random training/test partitions based on data collected in single-strain chips (no crosstalk data) to have an estimate of performance. Table 7 shows the average classification accuracy over the test set obtained across the 20 random partitions for each binary classification problem.

***Milestone 7.2:*** *Construction of a database of sensors' responses to chemicals of interest and null chemicals to be able to establish the statistical significance in detection:* Completed in Q4. We have constructed a database that is stored using MySQL in a Thecus NAS system capable of storing 22 TB of data. All of the image sequences gathered from the five device prototypes have been organized into a directory structure for each device and experiment. A master file containing the time stamp of each image and its file location has been compiled for quick image access via a Network File System (NFS). After being processed for rotation/translation and feature extraction, these images are the input variables to the machine learning algorithms. The training data is currently stored using the following columns of information:

1. Date time
2. Date time in milliseconds since Unix epoch
3. Experiment ID

| Toxin | Actual Conc. ($\mu$M) | Estimated Conc. ($\mu$M) | Std. Dev. of Est. Conc. | T-Test Result |
|---|---|---|---|---|
| Arsenic | 0.000000 | 0.094873 | 0.054253 | FAIL |
| | 0.100000 | 0.166705 | 0.039146 | PASS |
| | 0.200000 | 0.121520 | 0.045036 | PASS |
| | 0.500000 | 0.269255 | 0.042667 | PASS |
| | 1.000000 | 0.290522 | 0.029472 | PASS |
| Cadmium | 0.000000 | 0.073605 | 0.039364 | FAIL |
| | 0.025000 | 0.084992 | 0.035848 | PASS |
| | 0.050000 | 0.061783 | 0.019154 | FAIL |
| | 0.100000 | 0.081460 | 0.028670 | PASS |
| | 0.440000 | 0.258721 | 0.054637 | PASS |
| Chromium(VI) | 0.000000 | 0.402180 | 0.211879 | FAIL |
| | 0.200000 | 0.458967 | 0.162400 | PASS |
| | 0.500000 | 0.558889 | 0.163777 | PASS |
| | 1.000000 | 0.816885 | 0.133032 | PASS |
| | 5.000000 | 1.215800 | 0.324680 | PASS |
| Copper | 0.000000 | 1.037426 | 0.628584 | FAIL |
| | 0.100000 | 0.138169 | 0.120940 | FAIL |
| | 0.500000 | 0.890369 | 0.184934 | FAIL |
| | 1.000000 | 1.428817 | 0.437622 | PASS |
| | 2.000000 | 2.332876 | 0.320600 | PASS |
| | 5.000000 | 1.875258 | 0.285088 | PASS |
| | 10.000000 | 2.931468 | 0.816670 | PASS |
| Lead | 0.000000 | 0.674518 | 0.526677 | FAIL |
| | 0.600000 | 0.917133 | 0.294116 | PASS |
| | 1.800000 | 1.636931 | 0.326718 | PASS |
| | 3.600000 | 2.284836 | 0.987410 | PASS |
| | 7.200000 | 3.417676 | 1.085822 | PASS |
| Mercury | 0.000000 | 0.247743 | 0.185409 | FAIL |
| | 0.100000 | 0.241039 | 0.102734 | FAIL |
| | 0.200000 | 0.245659 | 0.087578 | FAIL |
| | 0.500000 | 0.392381 | 0.067889 | PASS |
| | 2.000000 | 0.968285 | 0.084330 | PASS |
| | 6.000000 | 1.548434 | 0.286386 | PASS |

*Table 8: Toxin detection results for each toxin at various concentrations using a two-sample T-test at a 0.05 significance level. The first column is the actual concentration, the second column is the average estimated concentration, the third column is the standard deviation of the estimation, and the fourth column is the result of the test.*

4. Device number
5. Toxin concentration at the start point
6. Concentration units at the start point
7. Toxin at the start point
8. Toxin at the start point using machine learning label (see below)
9. Multiple toxin flag at the start point (are there

multiple toxins? Y/N)
10. Toxin concentration at the end point
11. Concentration units at the end point
12. Toxin at the end point
13. Toxin at the end point using machine learning label
14. Multiple toxin flag at the end point
15. Responding strains
16. Additional notes
17. Delay time (time before the toxin(s) reach(es) the cells)
18. Success flag (did experiment complete successfully? Y/N)

The machine learning toxin labels are as follows: MilliQ_ddH20 = Code 1; Arsenic, As = 2; Cadmium, Cd = 3; Cobalt, Co = 4; Chromium(VI), Cr = 5; Copper, Cu = 6; Mercury, Hg = 7; Malathion, Mal = 8; Lead, Pb = 9; gfp_tracer = 10; and Ammonium, NH4-N = 11.

***Milestone 7.3:*** *Quantification of the GFP threshold of detection with respect to the concentration levels of the toxins):* Completed in Q3, with updated results through Q4. Based on the results of Milestone 7.1 (Table 7) and the ROC curves (Fig. 26), the models establish the limits of detection that are described in Table 8. Toxin detection results at various concentrations were determined by using a two-sample T-test at a 0.05 significance level. A nonlinear support vector regressor was used with the metaparameters cross-validated using 5-fold cross-validation.

***Milestone 7.4:*** *Construction of the Receiver Operating Characteristic curve (ROC) for the sensors to minimize false negatives and false positives):* Completed in Q3. The results in Milestone 7.1 were obtained by assuming that the penalties of misclassification are identical for positive and negative classes. In other words, the cost of classifying a GFP signal as "toxic" when it is not (or vice versa) is the same. However, it may be the case that the cost is not symmetric for positive and negative cases. A water sensor is a good example of this situation, since

it might be preferable to ensure high accuracy when toxins are actually in the water (true positive rate, TP) in exchange for increasing the number of cases that are classified as "toxin present" when there is not any toxin in the water (false positive rate, FP). The Receiver Operating Characteristic (ROC) curve is a 2-D parametrized curve used to quantify and represent the trade-off between the true positive rate and the false positive rate of a given classifier. The abscissa represents the False Positive rate, while the ordinate shows the True Positive rate. Therefore, the optimal classifier is represented by a point in the upper left corner of the ROC curve, since this point corresponds to the best possible case in which the classifier is able to correctly identify 100% of positive cases (toxin present) with no false alarms.

The parameter that defines the ROC curve in our classification model is the decision threshold, which determines whether a pattern (GFP signal) is classified as positive (toxin present) or negative (toxin not present). The SVM model provides a value (decision function) for each pattern that represents the confidence of the model in its prediction, and the final classification is obtained by assigning to the negative class those points with decision functions that are below the decision threshold, and classifying as positive samples those patterns with decision functions above this threshold. Therefore, by sweeping a grid of possible values for the SVM decision threshold, we obtained the ROC curves for the various toxins shown in Figure 26.

**Milestone 8: Use the models obtained in Milestone 7 to investigate whether the combination of nonspecific sensor responses to some toxins can be utilized to improve specificity**

*Summary:* We have successfully completed this milestone. We have demon-
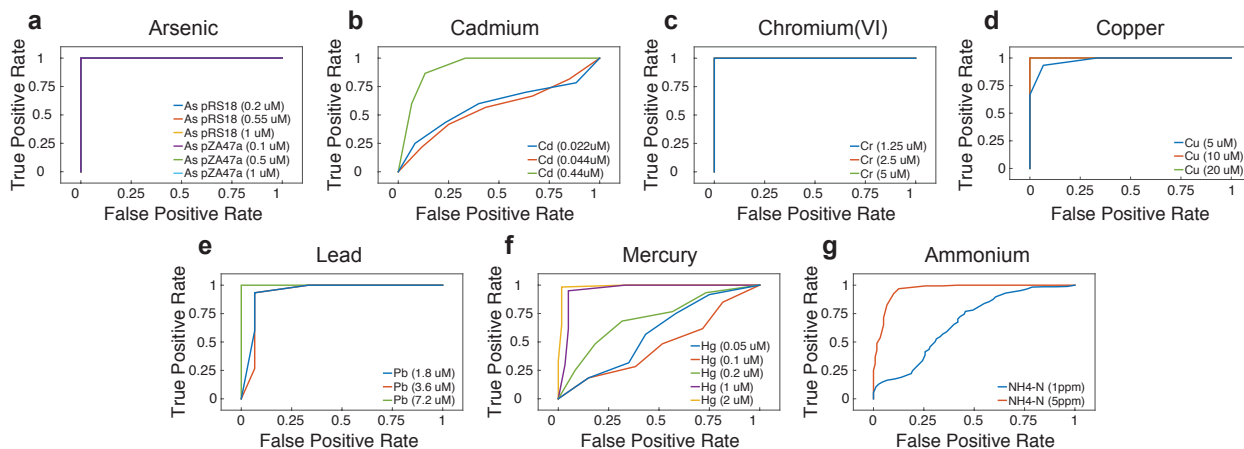
*Figure 26: Receiver Operating Characteristic (ROC) curves for the sensors. This parametrized curve quantifies the trade-off between the true positive rate and the false positive rate of a classifier. An optimal classifier is represented by a step function (e.g. arsenic), since this point corresponds to correct identification of 100% of the positive cases (toxin present) with no false alarms. In the other limit, worst-case classifiers are characterized by a diagonal ROC curve. Note that the arsenic, chromium(VI), copper, and lead sensors are nearly perfect at all tested levels, while very low levels of cadmium, mercury, and ammonium lead to suboptimal classification due to the detection limit of the genetic circuitry.*

strated the high specificity of the strains we have developed (Milestone 8.1), and we have used state-of-the-art pattern recognition algorithms to improve the performance of the classifier (Milestone 8.2).

***Milestone 8.1:*** *Estimation of the number of nonspecific sensors required to achieve maximum specificity in the discrimination of the target chemicals:* Completed in Q4. Because we used RNA-Seq to find promoters that were highly chemically specific, we found there to be only a small amount of crosstalk. The combinatoric information provided by non-specific responses was used to strengthen classification performance. To demonstrate this, we trained a classifier using a non-linear support vector machine on the GFP images concatenated to the estimations of the pixel derivatives for two different time scales. In Table 9 we present the confusion matrix of the sensor-specific strains to six different chemicals at various concentrations. The derivation of this matrix includes all of the available concentrations in the experiments, taking into account any non-specific responses of the sensor-specific

strains. Importantly, it demonstrates that the classifier can discriminate between the toxins. That is, in each row, there is one maximum number (*i.e.* the sensor responding to the correct toxin), and the rest of the numbers are uniformly low.

Generally, multi-class classification results nearing 80% represent a highly successful algorithm, particularly given the novelty of the devices and the data acquisition protocols. One issue contributing to negative results is the way in which we probed the sensors. We initially exposed the sensors to higher levels of toxins to verify that we could detect responses. However, once we established that we could sense these toxins well, we dropped down to much lower levels near the detection limit for the majority of the subsequent inductions. This weighted our data heavily toward the realm of low responses and low signal, skewing our results to contain more errors. As we have only begun to probe the parameter space, we expect to strengthen the confusion matrix significantly as we move forward.

We have performed a similar analysis of

|           | None  | Arsenic | Cadmium | Chromium(VI) | Copper | Lead  | Mercury |
|-----------|-------|---------|---------|--------------|--------|-------|---------|
| None      | 87.4% | 2.0%    | 2.6%    | 1.6%         | 0.8%   | 2.6%  | 3.0%    |
| Arsenic   | 6.2%  | 47.3%   | 7.4%    | 10.9%        | 9.3%   | 4.7%  | 14.3%   |
| Cadmium   | 6.7%  | 6.2%    | 73.3%   | 1.3%         | 4.6%   | 2.8%  | 5.1%    |
| Chromium(VI) | 4.8% | 10.6% | 1.5%    | 74.9%        | 3.4%   | 3.4%  | 1.3%    |
| Copper    | 2.1%  | 7.7%    | 4.5%    | 2.9%         | 76.5%  | 2.4%  | 3.9%    |
| Lead      | 6.4%  | 2.9%    | 2.6%    | 2.7%         | 2.3%   | 81.3% | 1.8%    |
| Mercury   | 6.2%  | 10.1%   | 4.0%    | 0.9%         | 3.1%   | 1.5%  | 74.2%   |

*Table 9: Confusion matrix solving a multiclass classification task for the following classes: baseline, Arsenic, Cadmiun, Chromium(VI), Copper, Lead, and Mercury. Cobalt also performs well, but there is not enough data to provide proper statistics. The machine learning algorithm used to calculate this table is a calibrated multiclass support vector machine using 10-fold cross validation on the data.[24] This matrix should be read from left to right. For example, lead is confused with baseline activity 6.4% of the time at low concentrations, while it rarely gets confused with Mercury. The main diagonal indicates how often the algorithms provide the right answer (e.g. 81.3% of the time for lead).*

our recent data collected using the 18-strain chip. This chip contains the new ammonium strain, which provides for a new confusion matrix, as seen in Table 10. This is a confusion matrix generated by taking the average of ten non-overlapping partitions of the training set (80% of the data) and the test set (20% of the data). Every training set of a partition was cross-validated to obtain the optimal metaparameters. After the cross-validation, the models were applied to the test set. We find that ammonium can be discriminated from the rest of the toxins 83% of the time.

***Milestone 8.2:*** *Apply state of the art pattern recognition algorithms using the database built for Milestone 7 to improve performance of the detection:* Completed in Q4. We have employed nonlinear support vector machine classification[24] on images that have undergone two types of transformation: 1) rotation and translation of the image with respect to a template for each of the sensor device setups (Fig. 27), and 2) an exponential moving average on the derivatives of the images to filter the pixel noise.[25,26]

We tried two different algorithms for image rotation and translation: minimization of the scalar product of the image with the template reference image, and a nor-
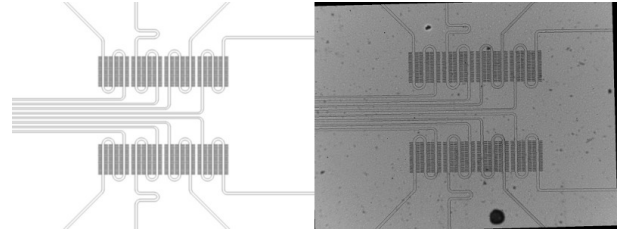


*Figure 27: Left: the template image used to align the images from all of the devices. Right: the result of applying the algorithm to an image.*

malized Euclidean distance from the bright field image to the template. We found the Euclidean distance to be the most effective method. The algorithm can run in real time to track the variations of the image's location in the device (Fig. 27). Once the images have been aligned, we can apply a feature extraction algorithm to zoom in on the traps where the strains are located.

The feature extraction algorithm involves two phases: image reduction and derivative calculation. The image reduction phase applies a smoothing algorithm on neighboring pixels according to a circular kernel using a radius of 41 pixels. The weights provided to the kernel have the maximum value in the center of the circle and decrease to a minimum value at the border of the circle. After the smoothing has been applied, the image is scaled from

|            | None   | Arsenic | Cadmium | Chromium(VI) | Cobalt | Copper | Lead   | Ammonium |
|------------|--------|---------|---------|--------------|--------|--------|--------|----------|
| None       | 45.02% | 7.66%   | 0.75%   | 13.48%       | 9.69%  | 2.56%  | 5.26%  | 15.64%   |
| Arsenic    | 3.50%  | 92.58%  | 0.00%   | 0.14%        | 0.00%  | 0.00%  | 0.00%  | 3.78%    |
| Cadmium    | 4.72%  | 0.00%   | 93.11%  | 2.17%        | 0.00%  | 0.00%  | 0.00%  | 0.00%    |
| Chromium(VI)| 8.05% | 0.18%   | 0.25%   | 90.43%       | 0.11%  | 0.11%  | 0.52%  | 0.35%    |
| Cobalt     | 9.75%  | 0.00%   | 0.00%   | 0.21%        | 89.32% | 0.49%  | 0.00%  | 0.22%    |
| Copper     | 15.73% | 0.00%   | 0.00%   | 1.12%        | 2.75%  | 78.27% | 0.00%  | 2.14%    |
| Lead       | 7.71%  | 0.00%   | 0.00%   | 1.22%        | 0.00%  | 0.00%  | 90.76% | 0.31%    |
| Ammonium   | 10.94% | 4.99%   | 0.00%   | 0.35%        | 0.12%  | 0.19%  | 0.12%  | 83.30%   |

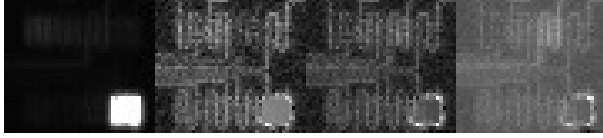*Table 10: Confusion matrix obtained from the 18-strain data.*



*Figure 28: An example of the results of the image transformation after applying the alignment, compression, and feature extraction on the GFP signal. This is the input signal that is provided to the classifier. The left square is the original image, and the other three squares show the changes made by the image processing techniques.*

$720\times640$ pixels to $45\times40$ pixels. This operation reduces the size of the feature space significantly.

The feature extraction algorithm next calculates the derivatives of the images to enhance the dynamic changes in the chip. The formula is calculated as $\Delta p_{ij}^k(t) = \alpha^k p_{ij}(t) + (1-\alpha)(1 - p_{ij}^k(t-1))$, and the time scales used in the filter are $\alpha^1 = 1/11$, $\alpha^2 = 1/26$, and $\alpha^3 = 1/51$. In the end, the total number of features is $(45\times40)\times4$. In Figure 28 we show the original GFP image and the result of the derivative estimation using three time scales $(\alpha^1, \alpha^2, \alpha^3)$.

After the feature extraction algorithm has been applied, the classifier is trained using 80% of the data and tested using the remaining 20%. We apply a 10-fold cross-validation on the 80% of the training data to determine the optimal model parameters before we test the model on the remaining 20%. The confusion matrices are then calculated by running an average on ten parti-

tions of train and test. For the 8-strain chip data, the cross-validated classification accuracy obtained is 76% with a standard deviation of 2.57%. The results are summarized in the confusion matrix (Table 9). Arsenic is the toxin that has the highest likelihood of being confused with other toxins. This is due to collecting a majority of our data at very low arsenic levels to test the limits of the sensor, which skews the results.

We have been continuously collecting data using five replicates of the sensor undergoing different alignment and feature extraction algorithms. Overall, we are extremely satisfied with the performance of this state-of-the-art classifier on this dataset, particularly given the limited amount of data collected thus far. Moreover, we have been intentionally probing the strains with concentration levels that are difficult to detect and classify. As we continue to collect data with all 18 strains, the classifier will improve further.

**Milestone 9: Develop the controller board to carry out sophisticated analysis of complex data**

*Summary:* We have successfully completed this milestone. A PandaBoard system on a chip (SoC) was selected for the primary control system which contains integrated Wi-Fi (Milestone 9.1). We have developed advanced image processing algorithms and embedded them on the PandaBoard (Milestone 9.2). We have developed

algorithms which can easily be stored and loaded on our PandaBoard SoC (Milestone 9.3).

***Milestone 9.1:*** *Design the controller board with wireless capability:* Completed in Q2. We tested two low-power platforms based on the Texas Instruments ARM processor with wireless capability: one based on the Sitara ARM and the other based on the Cortex-A9. We have opted for the more powerful Cortex-A9 due to the ease of use and low power consumption. We are currently using a PandaBoard, which is powered by a Texas Instruments OMAP4430 system on a chip (SoC) device. The OMAP4430 chipset contains a dual-core 1 GHz ARM Cortex-A9 MPCore CPU with 1 GB of DDR2 SDRAM, Wi-Fi capability, and an SD card slot offering up to 32 GB of storage. The electronics are similar to a modern smartphone in terms of processing power and power consumption. The PandaBoard solution allows us to install a Linux operating system so that we can use standard GNU compilers and run our software without any major modifications, which we demonstrate in the next two aims.

***Milestone 9.2:*** *Reduce the size of the pattern recognition algorithms to be able to be embedded in the PandaBoard.* Completed in Q4. Fluorescence images are used to train our classifier to be able to detect and discriminate between different toxins. In order to speed up the operation of the classifier, we must reduce its computational cost, which is directly linked to the number of images and the number of features in each image. Each image consists of a set of numerical features, each containing the intensity of a pixel in the image. Fortunately, the images contain many irrelevant features and regions as well as a large number of redundant features in neighboring pixels. Therefore, it is extremely useful to apply a feature selection algorithm to find the most infor-

mative features and to reduce the computational cost of the classifier. In Milestone 8.2, we describe the image processing and feature selection process that enabled us to reduce computational cost and embed our algorithms in the PandaBoard.

***Milestone 9.3:*** *Embed the algorithms in the microcontroller.* Completed in Q4. We have compiled and executed nonlinear support vector machines[24] on the images described in Milestone 9.2 after the feature selection process. We compiled the software, trained the model, and ran the trained model without any problems.

The algorithms were developed using the openMP API and compiled using GNU g++. Both are well established and stable options that run well on the PandaBoards. We implemented aggressive compiler optimizations to produce native code that runs fast on the multicore Cortex-A9 processors. Resulting performance was more than sufficient to allow real-time operation, with our embedded algorithms proving capable of classifying each image within 0.05 seconds.

## Updates for This Reporting Period

### *Meetings/Conferences/Workshops*

Hasty, J. (2014, June 4). Attended the DARPA Living Foundries meeting, Denver, CO.

Hasty, J., Huerta, R., & Tsimring, L. S. (2014, July 14). *An online biosensor for the protection of water supplies.* Presented at the Synthetic Biology: Engineering, Evolution & Design (SEED) conference, Manhattan Beach, CA.

### *Subject Invention Disclosures*

UCSD Docket No. SD2015-188: An Online Biosensor for the Protection of Water Supplies.
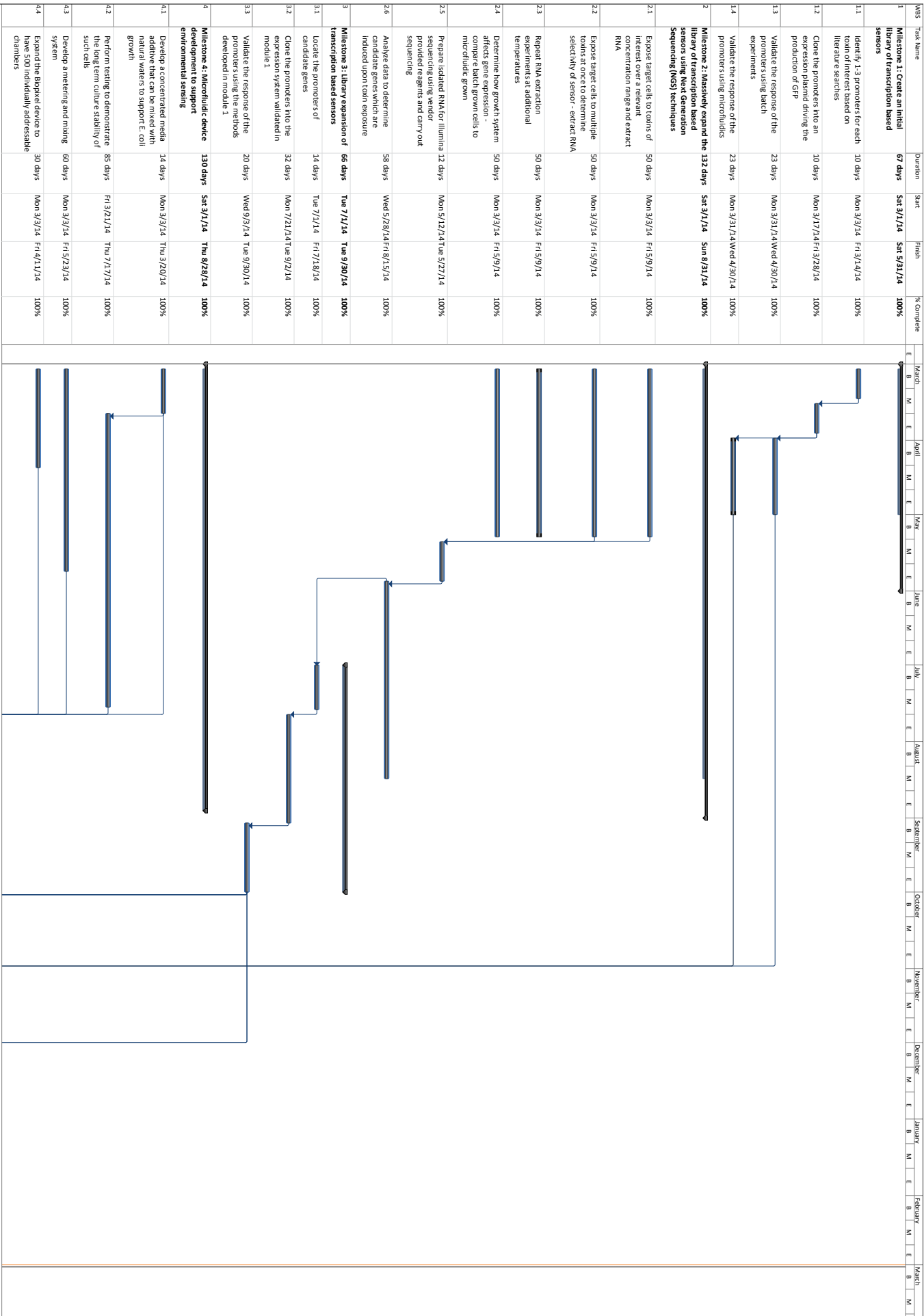
None at this time.

## Milestone Schedule Status

All tasks have been completed on schedule. Below find a Gantt chart showing the final task completion schedule.

| WBS | Task Name | Duration | Start | Finish | % Complete |
|---|---|---|---|---|---|
| 1 | **Milestone 1: Create an initial library of transcription based sensors** | **67 days** | **Sat 3/1/14** | **Sat 5/31/14** | **100%** |
| 1.1 | Identify 1-3 promoters for each toxin of interest based on literature searches | 10 days | Mon 3/3/14 | Fri 3/14/14 | 100% |
| 1.2 | Clone the promoters into an expression plasmid driving the production of GFP | 10 days | Mon 3/17/14 | Fri 3/28/14 | 100% |
| 1.3 | Validate the response of the promoters using batch experiments | 23 days | Mon 3/31/14 | Wed 4/30/14 | 100% |
| 1.4 | Validate the response of the promoters using microfluidics | 23 days | Mon 3/31/14 | Wed 4/30/14 | 100% |
| 2 | **Milestone 2: Massively expand the library of transcription based sensors using Next Generation Sequencing (NGS) techniques** | **132 days** | **Sat 3/1/14** | **Sun 8/31/14** | **100%** |
| 2.1 | Expose target cells to toxins of interest over a relevant concentration range and extract RNA | 50 days | Mon 3/3/14 | Fri 5/9/14 | 100% |
| 2.2 | Expose target cells to multiple toxins at once to determine selectivity of sensor - extract RNA | 50 days | Mon 3/3/14 | Fri 5/9/14 | 100% |
| 2.3 | Repeat RNA extraction experiments at additional temperatures | 50 days | Mon 3/3/14 | Fri 5/9/14 | 100% |
| 2.4 | Determine how growth system affects gene expression - compare batch grown cells to microfluidic grown | 50 days | Mon 3/3/14 | Fri 5/9/14 | 100% |
| 2.5 | Prepare isolated RNA for Illumina sequencing using vendor provided reagents and carry out sequencing | 12 days | Mon 5/12/14 | Tue 5/27/14 | 100% |
| 2.6 | Analyze data to determine candidate genes which are induced upon toxin exposure | 58 days | Wed 5/28/14 | Fri 8/15/14 | 100% |
| 3 | **Milestone 3: Library expansion of transcription based sensors** | **66 days** | **Tue 7/1/14** | **Tue 9/30/14** | **100%** |
| 3.1 | Locate the promoters of candidate genes | 14 days | Tue 7/1/14 | Fri 7/18/14 | 100% |
| 3.2 | Clone the promoters into the expression system validated in module-1 | 32 days | Mon 7/21/14 | Tue 9/2/14 | 100% |
| 3.3 | Validate the response of the promoters using the methods developed in module 1 | 20 days | Wed 9/3/14 | Tue 9/30/14 | 100% |
| 4 | **Milestone 4: Microfluidic device development to support environmental sensing** | **130 days** | **Sat 3/1/14** | **Thu 8/28/14** | **100%** |
| 4.1 | Develop a concentrated media additive that can be mixed with natural waters to support E. coli growth | 14 days | Mon 3/3/14 | Thu 3/20/14 | 100% |
| 4.2 | Perform testing to demonstrate the long term culture stability of such cells | 85 days | Fri 3/21/14 | Thu 7/17/14 | 100% |
| 4.3 | Develop a metering and mixing system | 60 days | Mon 3/3/14 | Fri 5/23/14 | 100% |
| 4.4 | Expand the Biopixel device to have 500 individually addressable chambers | 30 days | Mon 3/3/14 | Fri 4/11/14 | 100% |

| | | |
|---|---|---|
| Task | Milestone ◆ | Project Summary | External Milestone ◆ | Inactive Milestone ◆ | Manual Task | Start-only ⊏ | Deadline ◆ |
| Split | Summary | External Tasks | Inactive Task | Inactive Summary | Duration-only | Manual Summary Rollup | Finish-only ⊐ | Progress |

| WBS | Task Name | Duration | Start | Finish | % Complete |
|---|---|---|---|---|---|
| 4.5 | Increase the height of the Biopixel device's chambers to increase the output optical signal. Ensure the cell growth dynamics are unchanged | 75 days | Mon 3/3/14 | Fri 6/13/14 | |
| 4.6 | Develop a UV LED system that kills the cells as they emerge from the trapping region and enter waste collection | 60 days | Mon 3/3/14 | Fri 5/23/14 | 100% |
| 4.7 | Develop a method to freeze dry cells allowing them to be rehydrated with little loss of viability | 100 days | Mon 3/3/14 | Fri 7/18/14 | 100% |
| 4.8 | Develop a deposition technique to place cells into a region of a microfluidic device that is then bonded to a glass coverslip | 100 days | Mon 3/3/14 | Fri 7/18/14 | 100% |
| 4.9 | Develop and test microfluidic device in a laboratory environment | 29 days | Mon 7/21/14 | Thu 8/28/14 | 100% |
| 5 | Milestone 5: Low cost optical methods development | 65 days | Fri 8/1/14 | Thu 10/30/14 | 100% |
| 5.1 | Obtain an LED array for the GFP excitation of the cells in the microfluidic device | 35 days | Fri 8/1/14 | Thu 9/18/14 | 100% |
| 5.2 | Obtain optical filters for GFP excitation and emission which cover the entire area of the microfluidic device | 20 days | Fri 8/1/14 | Thu 8/28/14 | 100% |
| 5.3 | Obtain and characterize the performance of a low cost camera system to image the fluorescence signal of the device | 30 days | Fri 9/19/14 | Thu 10/30/14 | 100% |
| 5.4 | Replace the GFP fluorescence system with a luminescent system based on the Lux operon of V. fischeri | 30 days | Fri 8/1/14 | Thu 9/11/14 | 100% |
| 6 | Milestone 6: Device prototype development | 261 days | Sat 3/1/14 | Fri 2/27/15 | 100% |
| 6.1 | Find and characterize a low power pumping system that can pump natural water through the microfluidic device at a rate of 1 ml/hr | 40 days | Mon 3/3/14 | Fri 4/25/14 | 100% |
| 6.2 | Develop a filter system to prevent clogging of the microfluidic device | 80 days | Mon 3/3/14 | Fri 6/20/14 | 100% |
| 6.3 | Develop a flow reversal pumping regime to help clear the filter of contaminants | 40 days | Mon 6/23/14 | Fri 8/15/14 | 100% |
| 6.4 | Develop a waterproof enclosure that can be used to house the device and the electronics | 30 days | Mon 3/3/14 | Fri 4/11/14 | 100% |
| 6.5 | Develop a heating system to ensure the microfluidic device maintains the appropriate growth temperature | 60 days | Mon 3/3/14 | Fri 5/23/14 | 100% |
| 6.6 | Develop the electronics / software to coordinate the pumping regimes, image capture, data transmission and device power management | 80 days | Mon 3/3/14 | Fri 6/20/14 | 100% |

Project: Darpa Project Tasks 2014
Date: Sun 3/1/15

Task   Split   Milestone   Summary   Project Summary   External Tasks   External Milestone   Inactive Task   Inactive Milestone   Inactive Summary   Manual Task   Duration-only   Manual Summary Rollup   Manual Summary   Start-only   Finish-only   Deadline   Progress

44

| WBS | Task Name | Duration | Start | Finish | % Complete |
|---|---|---|---|---|---|
| 6.7 | Develop a chemical control stock to test the proper operation of the biosensor (ie add low doses of toxins for positive control) | 80 days | Mon 3/3/14 | Fri 6/20/14 | 100% |
| 6.8 | Develop a solar powered version of the prototype, including battery panels and charge controller | 40 days | Mon 6/23/14 | Fri 8/15/14 | 100% |
| 6.9 | Develop a completed device prototype and test in an outdoor environment | 80 days | Fri 10/31/14 | Thu 2/19/15 | 100% |
| 7 | Milestone 7: Build computational models to determine the threshold of detection for specific sensors based on models of experimental GFP responses | 126 days | Tue 7/1/14 | Tue 12/23/14 | 100% |
| 7.1 | Characterization of the family of GFP sensor responses to the set of chemicals of interest via computational models | 40 days | Wed 10/1/14 | Tue 11/25/14 | 100% |
| 7.2 | Construction of a database of sensors responses to chemicals of interest and null chemicals to be able to establish the statistical significance in detection | 60 days | Wed 10/1/14 | Tue 12/23/14 | 100% |
| 7.3 | Quantification of the GFP threshold of detection respect to the concentration levels of the toxins | 20 days | Wed 10/1/14 | Tue 10/28/14 | 100% |
| 7.4 | Construction of the Receiver operating characteristic curve (ROC) for the built sensors to minimize the false negatives and positives | 20 days | Wed 10/1/14 | Tue 10/28/14 | 100% |
| 8 | Milestone 8: Use the models obtained in Task 7 to investigate whether the combination of nonspecific sensor responses to some toxins can be utilized to improve specificity | 43 days | Mon 12/1/14 | Wed 1/28/15 | 100% |
| 8.1 | Estimation of the number of nonspecific sensors required to achieve maximum specificity in the discrimination of the target chemicals | 43 days | Mon 12/1/14 | Wed 1/28/15 | 100% |
| 8.2 | Apply state-of-the-art pattern recognition algorithms using the database built on task 7 to improve performance of the detection | 43 days | Mon 12/1/14 | Wed 1/28/15 | 100% |
| 9 | Milestone 9: Develop the controller board to carry out sophisticated analysis of complex data | 175 days | Tue 7/1/14 | Sun 3/1/15 | 100% |
| 9.1 | Design of the controlled board with wireless capability using Texas Instruments microcontrollers | 60 days | Tue 7/1/14 | Mon 9/22/14 | 100% |
| 9.2 | Reduce the size of the pattern recognition algorithms to be able to be embedded in the microcontroller | 60 days | Tue 9/23/14 | Mon 12/15/14 | 100% |
| 9.3 | Embed the algorithms in the microcontroller | 55 days | Tue 12/16/14 | Sun 3/1/15 | 100% |

| | | | | | |
|---|---|---|---|---|---|
| Task | | Milestone | ◆ | Project Summary | |
| Split | | Summary | | External Tasks | |
| | | External Milestone | ◆ | Inactive Milestone | ◇ |
| | | Inactive Task | | Inactive Summary | |
| | | Manual Task | | Manual Summary Rollup | |
| | | Duration-only | | Manual Summary | |
| | | Start-only | E | Finish-only | ] |
| | | Deadline | ◆ | Progress | |

45

[1] C. Xu, W. Shi, and B. P. Rosen. The chromosomal arsr gene of escherichia coli encodes a trans-acting metalloregulatory protein. *J Biol Chem*, 271(5):2427–2432, Feb 1996.

[2] J. Wu and B. P. Rosen. Metalloregulated expression of the ars operon. *J Biol Chem*, 268(1):52–58, Jan 1993.

[3] Judith Stocker, Denisa Balluch, Monika Gsell, Hauke Harms, Jessika Feliciano, Sylvia Daunert, Khurseed A. Malik, and Jan Roelof van der Meer. Development of a set of simple bacterial biosensors for quantitative and rapid measurements of arsenite and arsenate in potable water. *Environ Sci Technol*, 37(20):4743–4750, Oct 2003.

[4] Laura S. Busenlehner, Mario A. Pennella, and David P. Giedroc. The smtb/arsr family of metalloregulatory transcriptional repressors: Structural insights into prokaryotic metal resistance. *FEMS Microbiol Rev*, 27(2-3):131–143, Jun 2003.

[5] E. A. Permina, A. E. Kazakov, O. V. Kalinina, and M. S. Gelfand. Comparative genomics of regulation of heavy metal resistance in eubacteria. *BMC Microbiol*, 6:49, 2006.

[6] G. Endo and S. Silver. Cadc, the transcriptional regulatory protein of the cadmium resistance system of staphylococcus aureus plasmid pi258. *J Bacteriol*, 177(15):4437–4441, Aug 1995.

[7] S. W. Lee, E. Glickmann, and D. A. Cooksey. Chromosomal locus for cadmium resistance in pseudomonas putida consisting of a cadmium-transporting atpase and a merr family response regulator. *Appl Environ Microbiol*, 67(4):1437–1444, Apr 2001.

[8] Rita Branco, Ana Paula Chung, Tatiana Johnston, Volkan Gurel, Paula Morais, and Anatoly Zhitkovich. The chromate-inducible chrbacf operon from the transposable element tnotchr confers resistance to chromium(vi) and superoxide. *J Bacteriol*, 190(21):6996–7003, Nov 2008.

[9] Rita Branco, Armando Cristóvão, and Paula V. Morais. Highly sensitive, highly specific whole-cell bioreporters for the detection of chromate in environmental samples. *PLoS One*, 8(1):e54005, 2013.

[10] Esther Aguilar-Barajas, Selene Jacobo-Arreola, Luis A. Verduzco-Rosas, Rafael Jiménez-Mejía, Martha I. Ramírez-Díaz, Adriana Julián-Sánchez, Héctor Riveros-Rosas, and Carlos Cervantes. An lrp-type transcriptional regulator controls expression of the bacillus subtilis chromate transporter. *Antonie Van Leeuwenhoek*, 104(6):941–948, Dec 2013.

[11] Christopher Rensing and Gregor Grass. Escherichia coli mechanisms of copper homeostasis in a changing environment. *FEMS Microbiol Rev*, 27(2-3):197–213, Jun 2003.

[12] Anna Jarosławiecka and Zofia Piotrowska-Seget. Lead resistance in micro-organisms. *Microbiology*, 160(Pt 1):12–25, Jan 2014.

[13] Jon L. Hobman, Daniel J. Julian, and Nigel L. Brown. Cysteine coordination of pb(ii) is involved in the pbrr-dependent activation of the lead-resistance promoter, ppbra, from cupriavidus metallidurans ch34. *BMC Microbiol*, 12:109, 2012.

[14] A. Ivask, K. Hakkila, and M. Virta. Detection of organomercurials with sensor bacteria. *Anal Chem*, 73(21):5168–5171, Nov 2001.

[15] G. Nucifora, L. Chu, S. Silver, and T. K. Misra. Mercury operon regulation by the merr gene of the organomercurial resistance system of plasmid pdu1358. *J Bacteriol*, 171(8):4241–4247, Aug 1989.

[16] S. J. Park, J. Wireman, and A. O. Summers. Genetic analysis of the tn21 mer operator-promoter. *J Bacteriol*, 174(7):2160–2171, Apr 1992.

[17] L. Chu, D. Mukhopadhyay, H. Yu, K. S. Kim, and T. K. Misra. Regulation of the staphylococcus aureus plasmid pi258 mercury resistance operon. *J Bacteriol*, 174(21):7044–7047, Nov 1992.

[18] D. Rother, R. Mattes, and J. Altenbuchner. Purification and characterization of merr, the regulator of the broad-spectrum mercury resistance genes in streptomyces lividans 1326. *Mol Gen Genet*, 262(1):154–162, Aug 1999.

[19] Dennis J. Paustenbach, Brent L. Finley, Fionna S. Mowat, and Brent D. Kerger. Human health risk and exposure assessment of chromium (vi) in tap water. *J Toxicol Environ Health A*, 66(14):1295–1339, Jul 2003.

[20] Matthew D. Stout, Ronald A. Herbert, Grace E. Kissling, Bradley J. Collins, Gregory S. Travlos, Kristine L. Witt, Ronald L. Melnick, Kamal M. Abdo, David E. Malarkey, and Michelle J. Hooth. Hexavalent chromium is carcinogenic to f344/n rats and b6c3f1 mice after chronic oral exposure. *Environ Health Perspect*, 117(5):716–722, May 2009.

[21] Tor Norseth. The carcinogenicity of chromium. *Environmental health perspectives*, 40:121, 1981.

[22] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[23] Richard O Duda, Peter E Hart, and David G Stork. *Pattern classification*. John Wiley & Sons, 2012.

[24] Ramón Huerta, Shankar Vembu, J Amigó, Thomas Nowotny, and Charles Elkan. Inhibition in multiclass classification. *Neural computation*, 24(9):2473–2507, 2012.

[25] Alexander Vergara, Jordi Fonollosa, Jonas Mahiques, Marco Trincavelli, Nikolai Rulkov, and Ramón Huerta. On the performance of gas sensor arrays in open sampling systems using inhibitory support vector machines. *Sensors and Actuators B: Chemical*, 185:462–477, 2013.

[26] Alexander Vergara, Shankar Vembu, Tuba Ayhan, Margaret A Ryan, Margie L Homer, and Ramón Huerta. Chemical gas sensor drift compensation using classifier ensembles. *Sensors and Actuators B: Chemical*, 166:320–329, 2012.

[27] Ben Langmead, Cole Trapnell, Mihai Pop, and Steven L. Salzberg. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*, 10(3):R25, 2009.

[28] Aaron R. Quinlan and Ira M. Hall. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6):841–842, Mar 2010.

[29] Simon Anders, Paul Theodor Pyl, and Wolfgang Huber. HTSeq–a Python framework to work with high-throughput sequencing data. *bioRxiv*, 2014.

[30] Simon Anders and Wolfgang Huber. Differential expression analysis for sequence count data. *Genome Biol*, 11(10):R106, 2010.

[31] Simon Anders, Davis J. McCarthy, Yunshun Chen, Michal Okoniewski, Gordon K. Smyth, Wolfgang Huber, and Mark D. Robinson. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat Protoc*, 8(9):1765–1786, Sep 2013.

***Milestone 1.2:*** *Clone the promoters into an expression plasmid driving the production of GFP:* Synthetic constructs were transformed into both standard *E. coli* MG1655 and optimized *E. coli* LABEC01 cells for expression analysis. By serially passing MG1655 in M9 medium through several generations, we evolved the common MG1655 lab strain, which has been adapted for growth in rich lysogeny broth (LB) medium, into a strain well-suited to growth in M9 minimal medium within our microfluidic devices. The M9-adapted strain, which we named LABEC01, exhibited multiple phenotypic changes. The growth rate in minimal medium increased by approximately 10%, colonies on minimal medium plates were observed to be smoother as compared to the parent MG1655, and the bacteria aggregated less when grown in microfluidic devices. To investigate how these adaptations affect the cellular response to toxins and the activity of sensitive promoters, we also exposed strain LABEC01 to the full set of toxins for RNA-Seq analysis.

# Candidate toxin-responsive promoters

| Toxin | Gene/ Promoter | Source | RBS | Host Strain/ Plasmid | Concentration Sensed in Microfluidic Device ($\mu$M) | SNR after 6 h |
|---|---|---|---|---|---|---|
| Ammonium | $p_{nasA}$ | *B. subtilis* genome | native | | | |
| Ammonium | $p_{nasA}$ | *B. subtilis* genome | synthetic | | | |
| Ammonium | $p_{nasB}$ | *B. subtilis* genome | native | | | |
| Ammonium | $p_{nasB}$ | *B. subtilis* genome | synthetic | | | |
| Ammonium | $p_{spo1-tnrA1}$ | *B. subtilis* genome | synthetic | | | |
| Ammonium | $p_{spo1-tnrA2}$ | *B. subtilis* genome | synthetic | | | |
| Arsenic | $arsR/p_{arsR}$ | *E. coli* plasmid | native | *E. coli* MG1655/As1 | 0.13 | 20 |
| Arsenic | $arsR/p_{arsR}$ | *E. coli* genome | synthetic | *E. coli* LABEC01/As3 | | |
| Arsenic | $arsR/p_{arsR}$ | *E. coli* genome | synthetic | *E. coli* MG1655/As3 | 0.13 | 33 |
| Arsenic | $arsR/p_{arsR}$ | *S. aureus* plasmid | native | *E. coli* MG1655/As5 | | |
| Arsenic | $p_{arsR}$ | *E. coli* RNA-Seq | | | | |
| Cadmium | $cadC/p_{cadC}$ | *S. aureus* plasmid | native | *E. coli* MG1655/Cd1 | 0.04 | 17 |
| Cadmium | $cadC/p_{cadC}$ | *S. aureus* plasmid | synthetic | *E. coli* MG1655/Cd2 | | |
| Cadmium | $cadR/p_{cadR}$ | *P. putida* genome | native | *E. coli* MG1655/Cd4 | | |
| Cadmium | $cadR/p_{cadR}$ | *P. putida* genome | synthetic | *E. coli* MG1655/Cd3 | | |
| Cadmium | $p_{zntA}$ | *E. coli* RNA-Seq | | | | |
| Chromium(VI) | $chrB/p_{chrB}$ | *C. metallidurans* plasmid | native | *E. coli* MG1655/Cr3 | | |
| Chromium(VI) | $chrB/p_{chrB}$ | *C. metallidurans* plasmid | synthetic | *E. coli* MG1655/Cr2 | | |
| Chromium(VI) | $chrB/p_{chrB}$ | *O. tritici* transposon | native | *E. coli* LABEC01/Cr5 | 5 | 5 |
| Chromium(VI) | $chrB/p_{chrB}$ | *O. tritici* transposon | native | *E. coli* MG1655/Cr5 | | |
| Chromium(VI) | $chrB/p_{chrB}$ | *O. tritici* transposon | synthetic | *E. coli* LABEC01/Cr4 | | |
| Chromium(VI) | $chrB/p_{chrB}$ | *O. tritici* transposon | synthetic | *E. coli* MG1655/Cr4 | | |
| Chromium(VI) | $chrS/p_{chrS}$ | *B. subtilis* genome | synthetic | *E. coli* MG1655/Cr1 | | |
| Chromium(VI) | $p_{recN}$ | *E. coli* RNA-Seq | | | | |
| Chromium(VI) | $p_{sulA}$ | *E. coli* RNA-Seq | | | | |
| Chromium(VI) | $p_{umuD}$ | *E. coli* RNA-Seq | | | | |
| Cobalt | $p_{dadA}$ | *E. coli* RNA-Seq | | | | |
| Cobalt | $p_{hmp}$ | *E. coli* RNA-Seq | | | | |
| Cobalt | $p_{ilvB}$ | *E. coli* RNA-Seq | | | | |
| Cobalt | $p_{lipA}$ | *E. coli* RNA-Seq | | | | |
| Cobalt | $p_{mmuP}$ | *E. coli* RNA-Seq | native | *E. coli* MG1655/Co7 | | |
| Cobalt | $p_{mmuP}$ | *E. coli* RNA-Seq | synthetic | *E. coli* MG1655/Co8 | | |
| Cobalt | $nmtR/p_{nmtR}$ | *M. tuberculosis* genome | native | *E. coli* MG1655/Co1 | | |
| Cobalt | $nmtR/p_{nmtR}$ | *M. tuberculosis* genome | synthetic | *E. coli* MG1655/Co2 | | |
| Cobalt | $p_{soxR}$ | *E. coli* RNA-Seq | | | | |
| Cobalt | $p_{tehA}$ | *E. coli* RNA-Seq | | | | |
| Cobalt | $p_{ygbA}$ | *E. coli* RNA-Seq | native | *E. coli* MG1655/Co3 | | |
| Cobalt | $p_{ygbA}$ | *E. coli* RNA-Seq | synthetic | *E. coli* MG1655/Co4 | | |
| Cobalt | $p_{yjbJ}$ | *E. coli* RNA-Seq | native | *E. coli* MG1655/Co5 | | |
| Cobalt | $p_{yjbJ}$ | *E. coli* RNA-Seq | synthetic | *E. coli* MG1655/Co6 | | |
| Cobalt | $p_{yqfA}$ | *E. coli* RNA-Seq | | | | |
| Copper | $cueR/p_{copA}$ | *E. coli* genome | native | *E. coli* MG1655/Cu1 | 25 | 65 |
| Copper | $(cusS/R)/p_{cusC}$ | *E. coli* genome | native | *E. coli* MG1655/Cu2 | | |
| Copper | $p_{cusC}$ | *E. coli* RNA-Seq | | | | |
| Copper | $p_{cusR}$ | *E. coli* RNA-Seq | | | | |

Table 11 – continued from previous page

| Toxin | Gene/Promoter | Source | RBS | Host Strain/Plasmid | Concentration Sensed in Microfluidic Device ($\mu$M) | SNR after 6 h |
|---|---|---|---|---|---|---|
| Lead | $pbrR$/$p_{pbrR}$ | *C. metallidurans* plasmid | native | *E. coli* LABEC01/Pb1 | | |
| Lead | $pbrR$/$p_{pbrR}$ | *C. metallidurans* plasmid | native | *E. coli* MG1655/Pb1 | | |
| Lead | $pbrR$/$p_{pbrR}$ | *C. metallidurans* plasmid | synthetic | *E. coli* LABEC01/Pb2 | 7 | 18 |
| Lead | $pbrR$/$p_{pbrR}$ | *C. metallidurans* plasmid | synthetic | *E. coli* MG1655/Pb2 | | |
| Lead | $p_{mntH}$ | *E. coli* RNA-Seq | | | | |
| Lead | $p_{shiA}$ | *E. coli* RNA-Seq | | | | |
| Lead | $p_{ybiI}$ | *E. coli* RNA-Seq | | | | |
| Lead | $p_{yjjZ}$ | *E. coli* RNA-Seq | | | | |
| Malathion | $p_{cusC}$ | *E. coli* RNA-Seq | | | | |
| Malathion | $p_{nemR}$ | *E. coli* RNA-Seq | | | | |
| Mercury | $merR$/$p_{merR}$ | *E. coli* plasmid | native | *E. coli* MG1655/Hg4 | | |
| Mercury | $merR$/$p_{merR}$ | *E. coli* plasmid | synthetic | *E. coli* MG1655/Hg3 | 0.1 | 20 |
| Mercury | $merR$/$p_{merR}$ | *S. aureus* plasmid | native | *E. coli* MG1655/Hg2 | | |
| Mercury | $merR$/$p_{merR}$ | *S. aureus* plasmid | synthetic | *E. coli* MG1655/Hg1 | | |
| Mercury | $merR$/$p_{merR}$ | *S. marcescens* plasmid | native | *E. coli* MG1655/Hg6 | | |
| Mercury | $merR$/$p_{merR}$ | *S. marcescens* plasmid | synthetic | *E. coli* MG1655/Hg5 | | |
| Zinc | $p_{mntH}$ | *E. coli* RNA-Seq | | | | |
| Zinc | $p_{shiA}$ | *E. coli* RNA-Seq | | | | |
| Zinc | $p_{yjjZ}$ | *E. coli* RNA-Seq | | | | |
| Zinc | $p_{zntA}$ | *E. coli* RNA-Seq | | | | |
| Zinc | $p_{zraP}$ | *E. coli* RNA-Seq | | | | |

*Table 11: All candidate toxin-responsive promoters identified in this work, ordered by the toxin of expected sensitivity. In the case of promoters identified by RNA-Seq, the gene is unknown. For promoters that have been expressed in a synthetic construct, the selected RBS and host strain are shown. If this synthetic construct has been used to sense the toxin within a microfluidic device, the concentration sensed and SNR after 6 h are shown.*

**RNA-Seq results for promoter activation in *E. coli* MG1655 in response to single and multiple toxin exposures at low and high concentrations:**

Our analysis of the sequencing data from RNA-Seq experiments to determine candidate genes that are induced upon toxin exposure consisted of three main tasks: sequence alignment, quantification of gene expression, and identification of differentially expressed genes.

**Sequence alignment:** Reads were aligned to the reference *E. coli* K-12 substr. MG1655 genome using a tolerance of at most two mismatches per alignment to protect against sequencing errors. The alignment was performed using Bowtie software,[27] which is known to be very efficient in aligning reads to a reference genome.

**Quantification of gene expression:** The expression level of each gene was determined as a function of the number of aligned reads mapping to the gene. After analyzing several approaches adopted in the literature to tabulate the number of reads mapping to each gene, we implemented our own software capable of reproducing the counting algorithms behind some of the standard toolboxes such as Bedtools[28] and HTSeq.[29] In particular, we counted the number of reads mapping to each gene regardless of whether the read mapped to several genes, taking into account the strand-specificity of each read. Additionally, we implemented our own algorithms for sequence alignment and quantification of gene expression in order to crosscheck all results.

**Identification of differentially expressed genes:** Finally, a set of statistical and information theory algorithms were applied in order to extract not only differentially expressed (DE) genes for each toxin with respect to the control samples (pure water) but also toxin-specific genes.

DESeq is a standard tool for identifying DE genes that allowed us to select sensitive genes with differential expression between the control samples (pure water) and the cells exposed to toxin. It assumes that the number of counts for each gene across experimental replicates follows a negative binomial distribution.[30,31] We considered genes with a False Discovery Rate (FDR) lower than 1% as DE in order to ensure statistically robust DE genes. We note that some genes showed high variability in the control samples across different batches of RNA-Seq experiments, indicating that these genes are very sensitive to environmental conditions. We identified 846 of these genes by performing a DESeq differential analysis (FDR<1%) between the control samples in different batches and subsequently removed them from the candidate pool. The number of DE genes (FDR<1%) identified for each condition when compared to the negative samples in the same batch and after removing genes that are DE between control samples is given in Table 6.

Ideally, good candidate specific genes are those with a significant fold-change with respect to the control samples but with a negligible fold-change with respect to the other toxins. Additionally, genes with the largest number of counts and expression levels are preferable in order to maximize the signal-to-noise ratio. When it is not possible to find toxin-specific genes, the next generation of good candidates is formed by those genes satisfying the above properties for a small subset of toxins (multiple-toxin response). It is desirable to have single-toxin-specific genes for several of the toxins in the combination in order to determine toxin-specific multi-gene-responses by means of logical operations.

In rare cases, shared genes are differentially expressed. Therefore, we have developed information theoretic measures to improve the toxin separability. The core idea of the approach is that low entropies (or highly informative genes) correspond to toxin-specific genes, while large entropies (low information) are associated with scenarios in which DE fold-changes across different

toxins are similar and should be discarded. The result of the analysis shows that toxins can easily be discriminated by using simple boolean rules as shown in the main report.
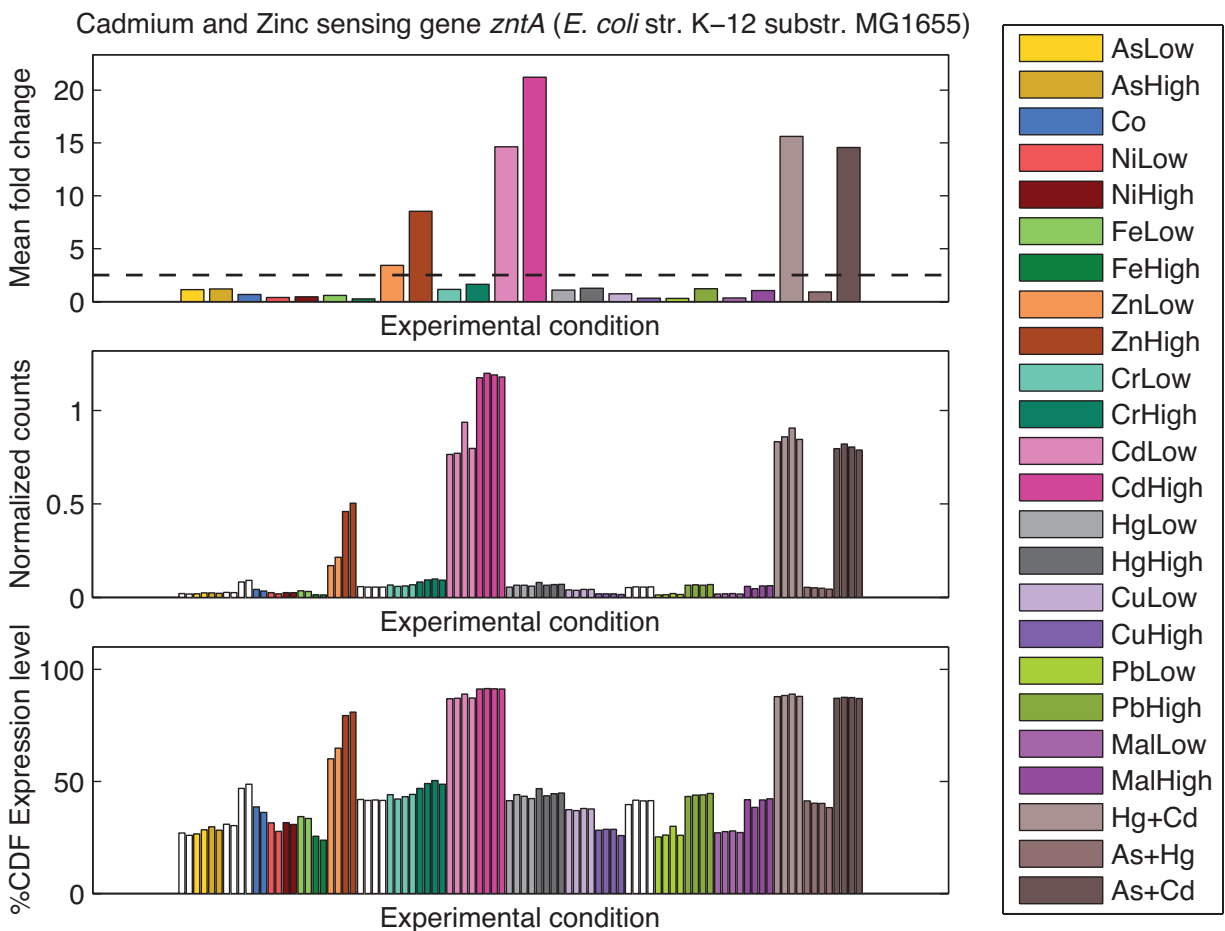


*Figure 29: Response of the zntA promoter in* **E. coli** *MG1655 for sensing cadmium and zinc. Note that the zntA promoter responds monotonically to increasing concentrations of cadmium alone and in combination with other heavy metals (rightmost conditions) without exhibiting crosstalk. A cadmium-specific sensor can be implemented by combining this response with the zinc-specific response of zraP using the boolean expression (zntA)∧(¬zraP).*
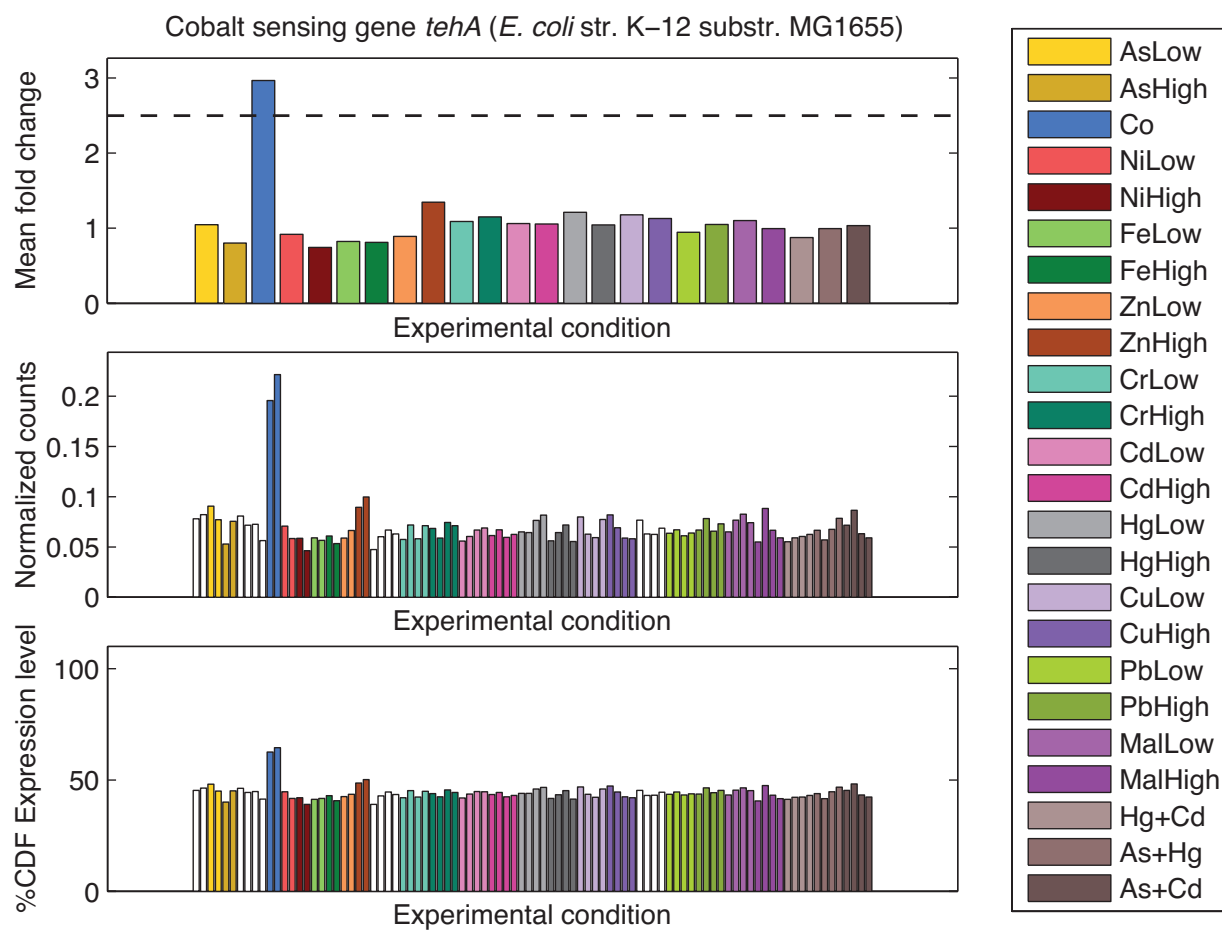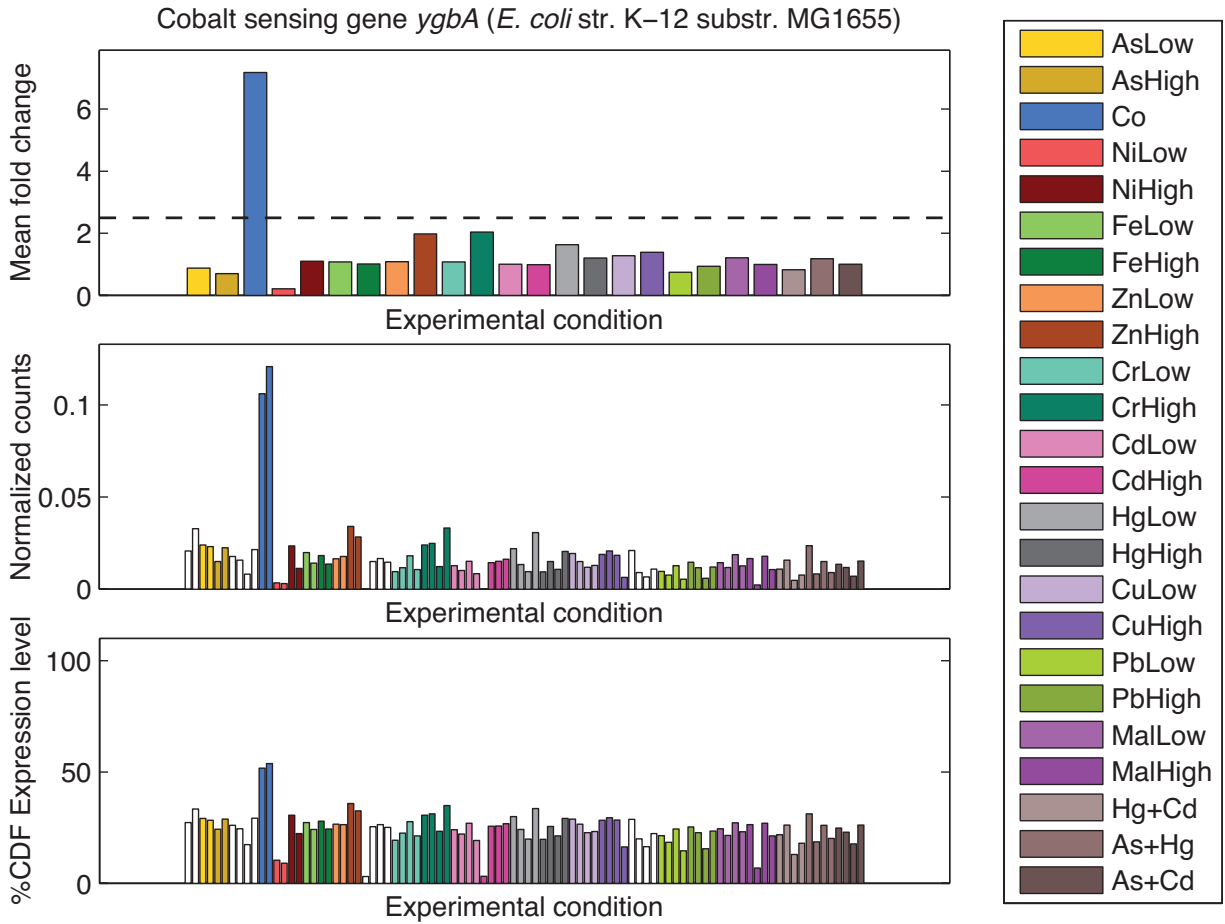
*Figure 30: Response of the recN promoter in* **E. coli** *MG1655 for specifically sensing chromium(VI). Note that the response of the recN promoter is specific to chromium(VI) alone but only shows sensitivity to high concentrations of chromium(VI).*
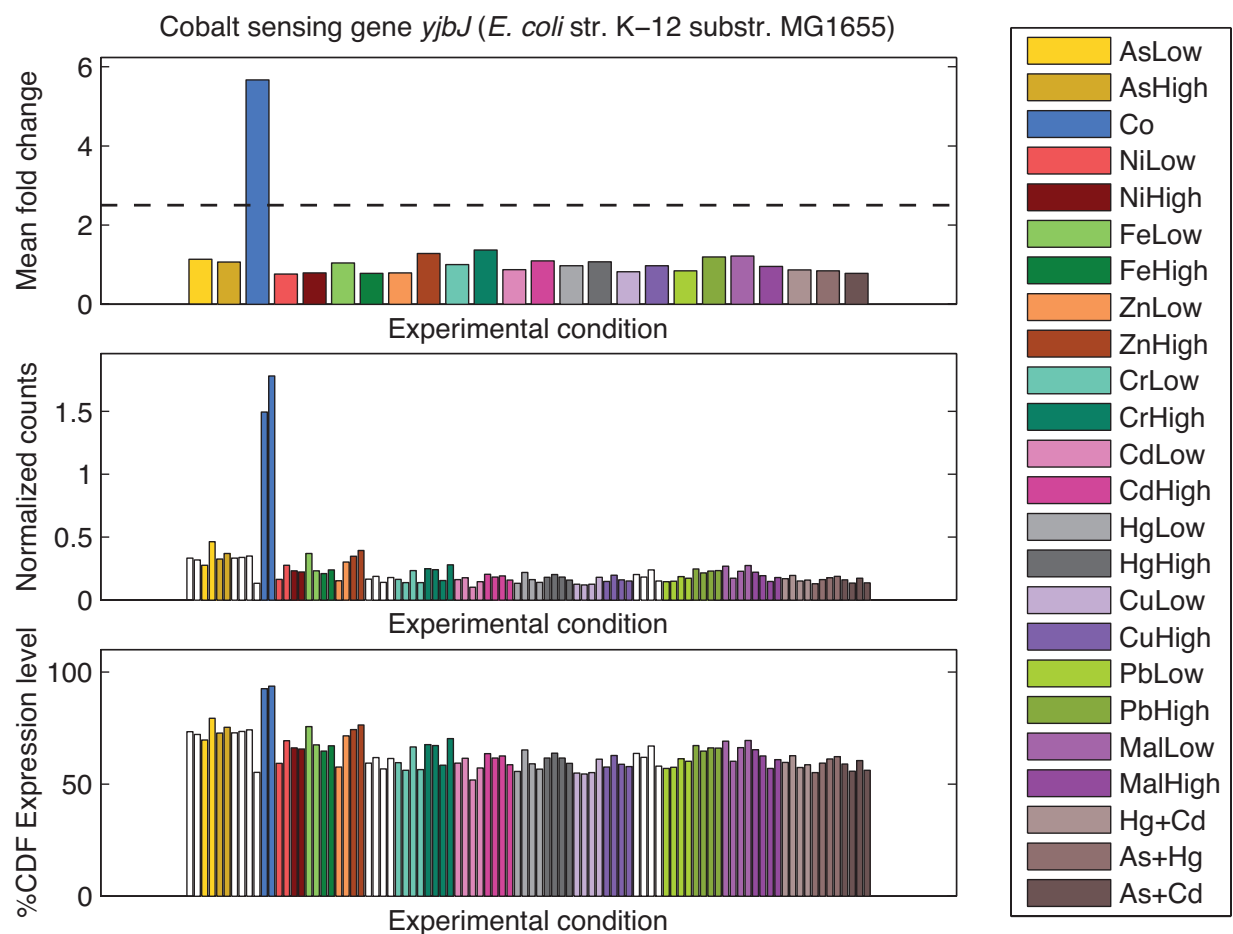
*Figure 31: Response of the sulA promoter in* **E. coli** *MG1655 for specifically sensing chromium(VI). Note that the response of the sulA promoter is specific to chromium(VI) alone but only shows sensitivity to high concentrations of chromium(VI).*
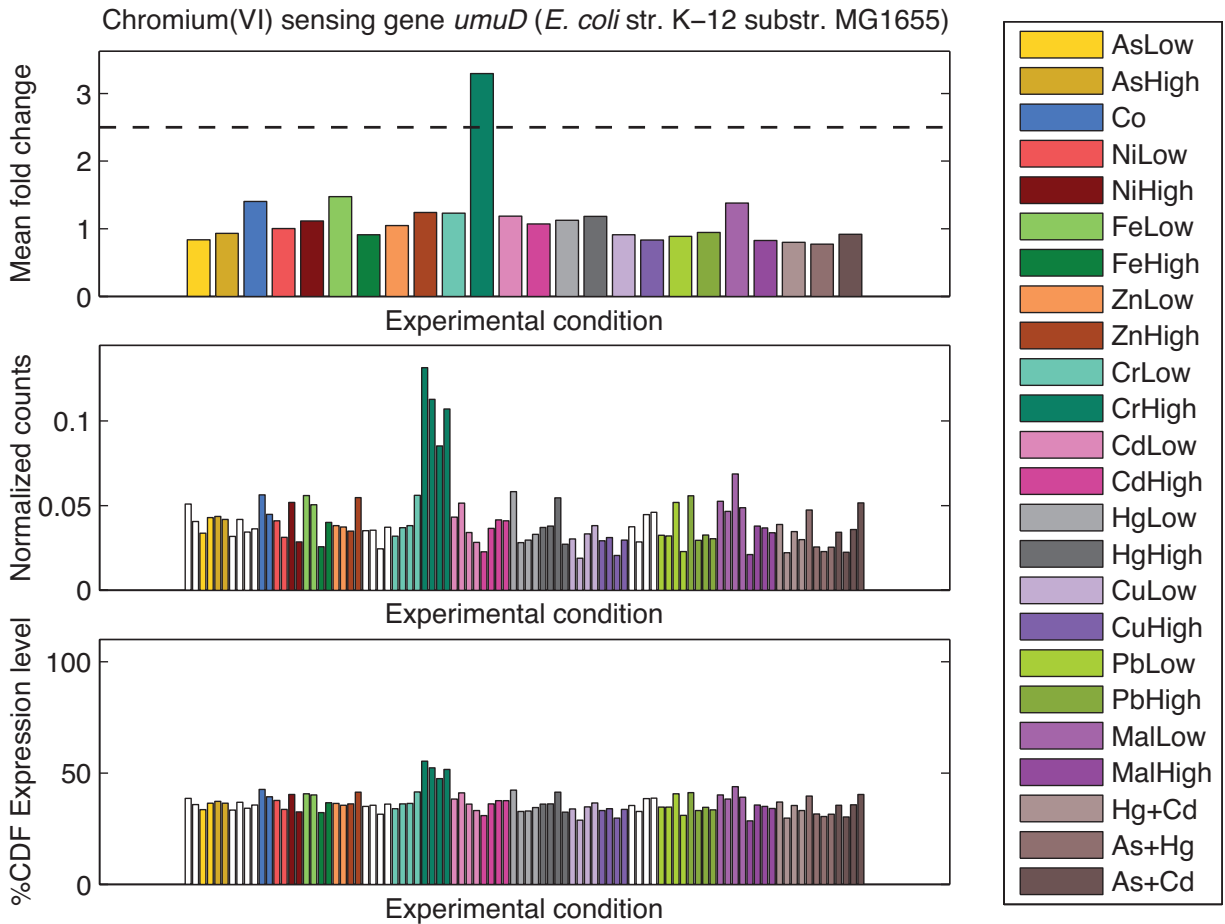
*Figure 32: Response of the umuD promoter in* **E. coli** *MG1655 for specifically sensing chromium(VI). Note that the response of the umuD promoter is specific to chromium(VI) alone but only shows sensitivity to high concentrations of chromium(VI).*
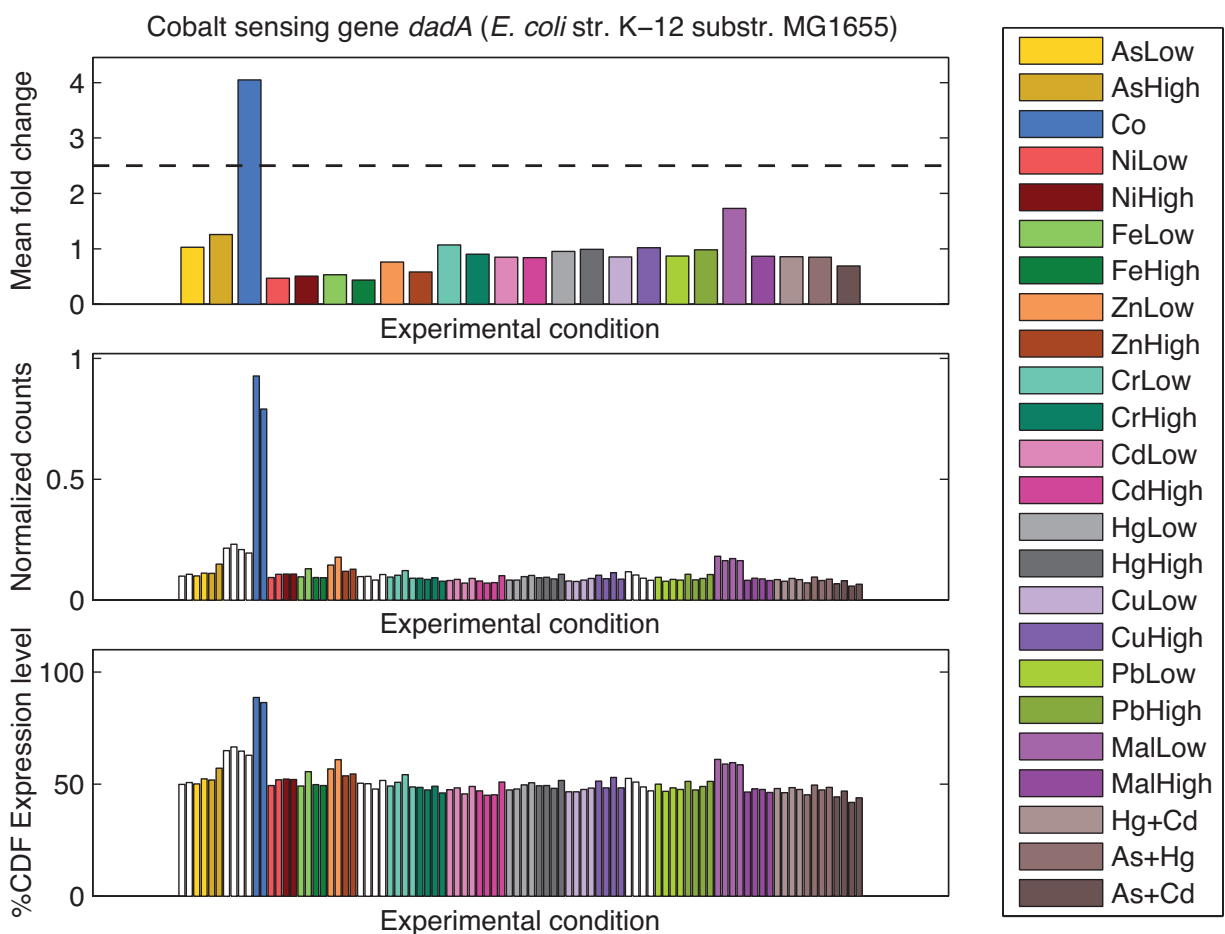
*Figure 33: Response of the dadA promoter in* **E. coli** *MG1655 for specifically sensing cobalt. Note that the response of the dadA promoter is specific to cobalt alone.*
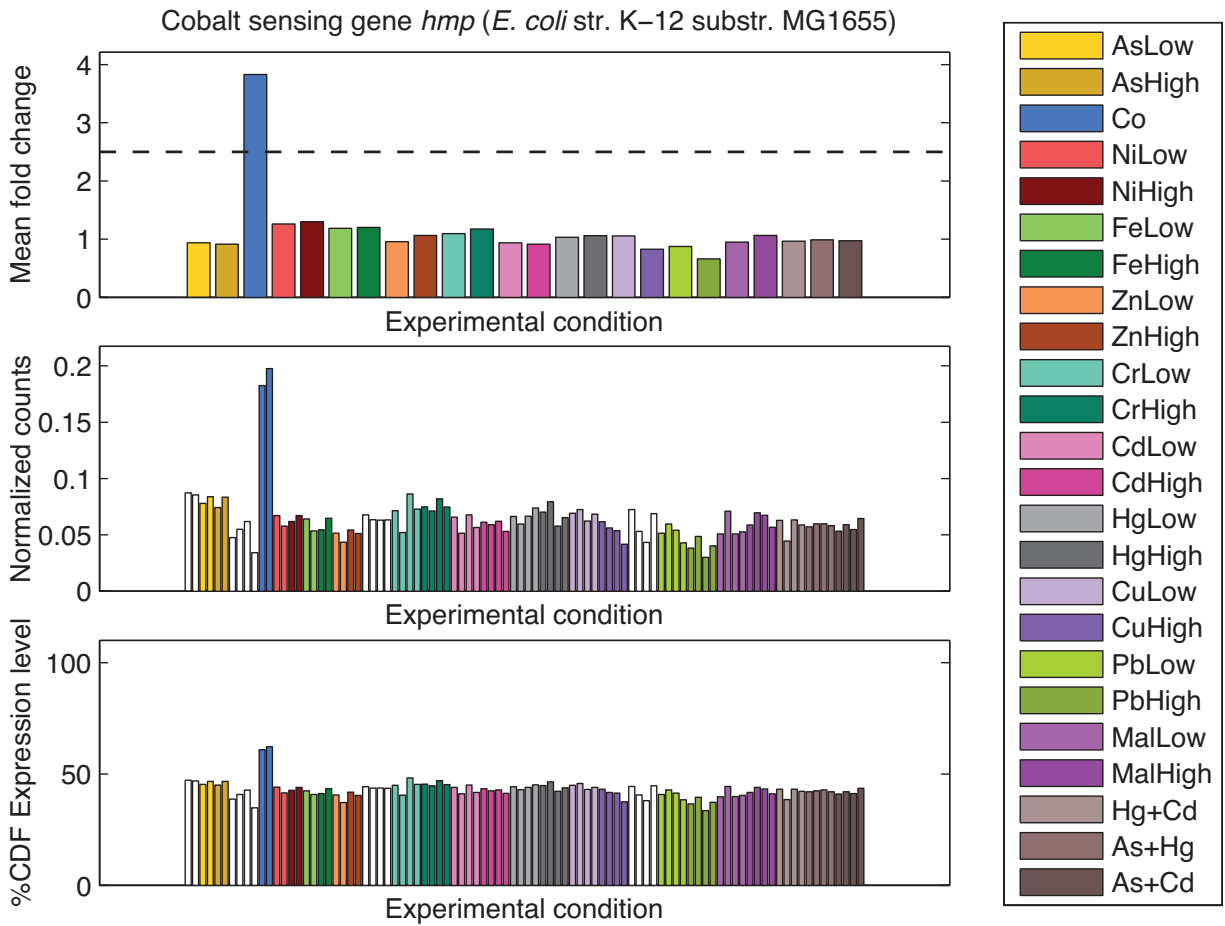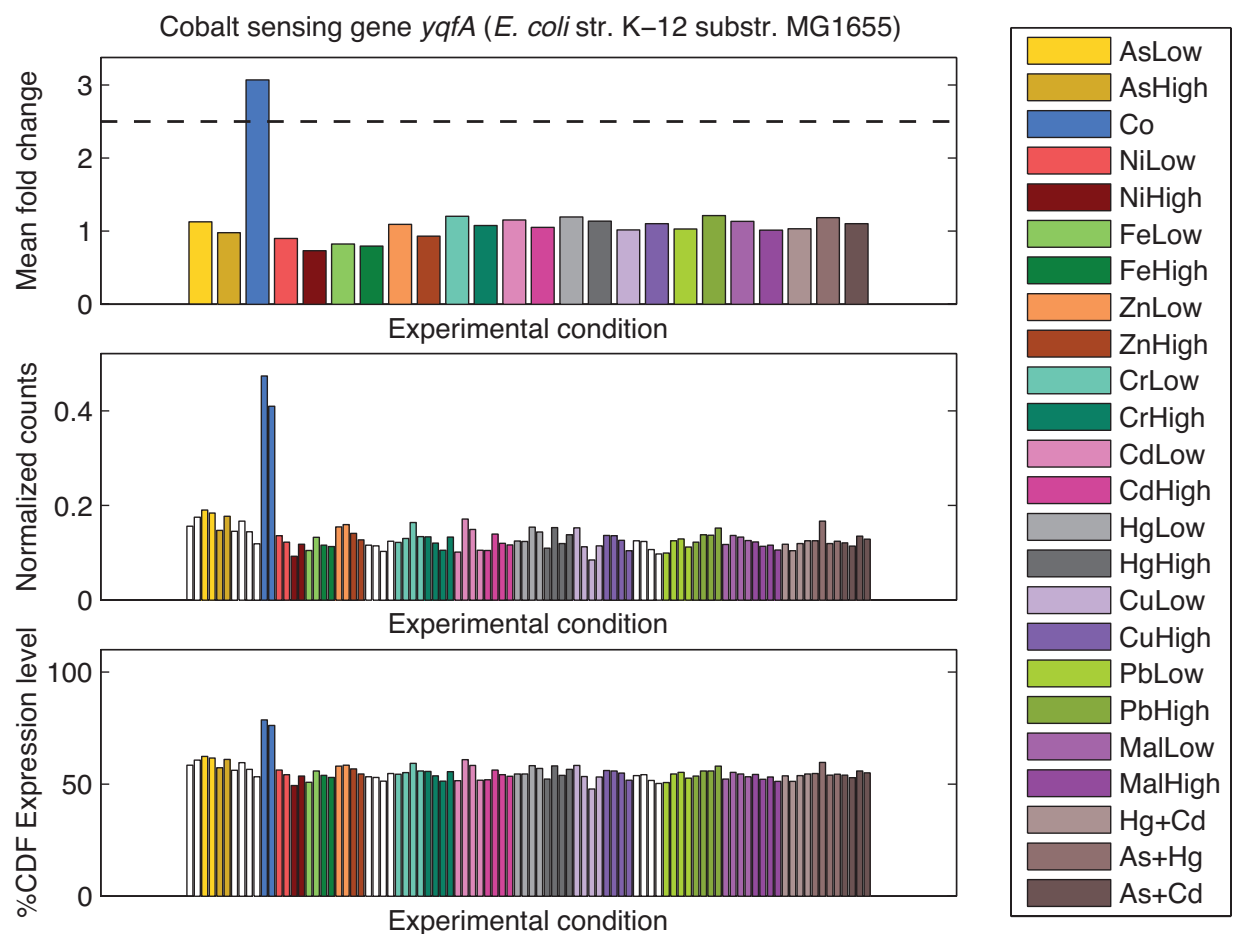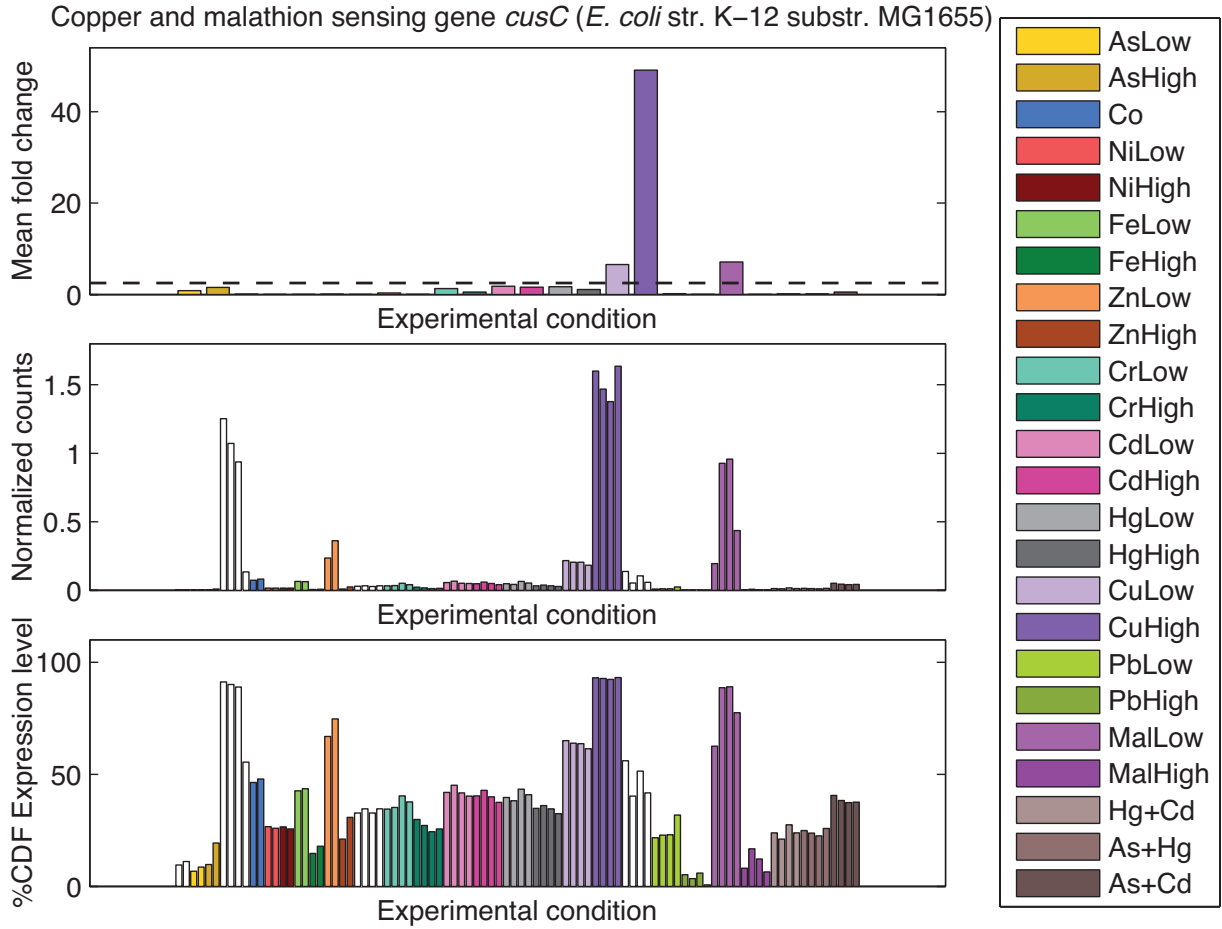
*Figure 34: Response of the hmp promoter in* **E. coli** *MG1655 for specifically sensing cobalt. Note that the response of the hmp promoter is specific to cobalt alone.*
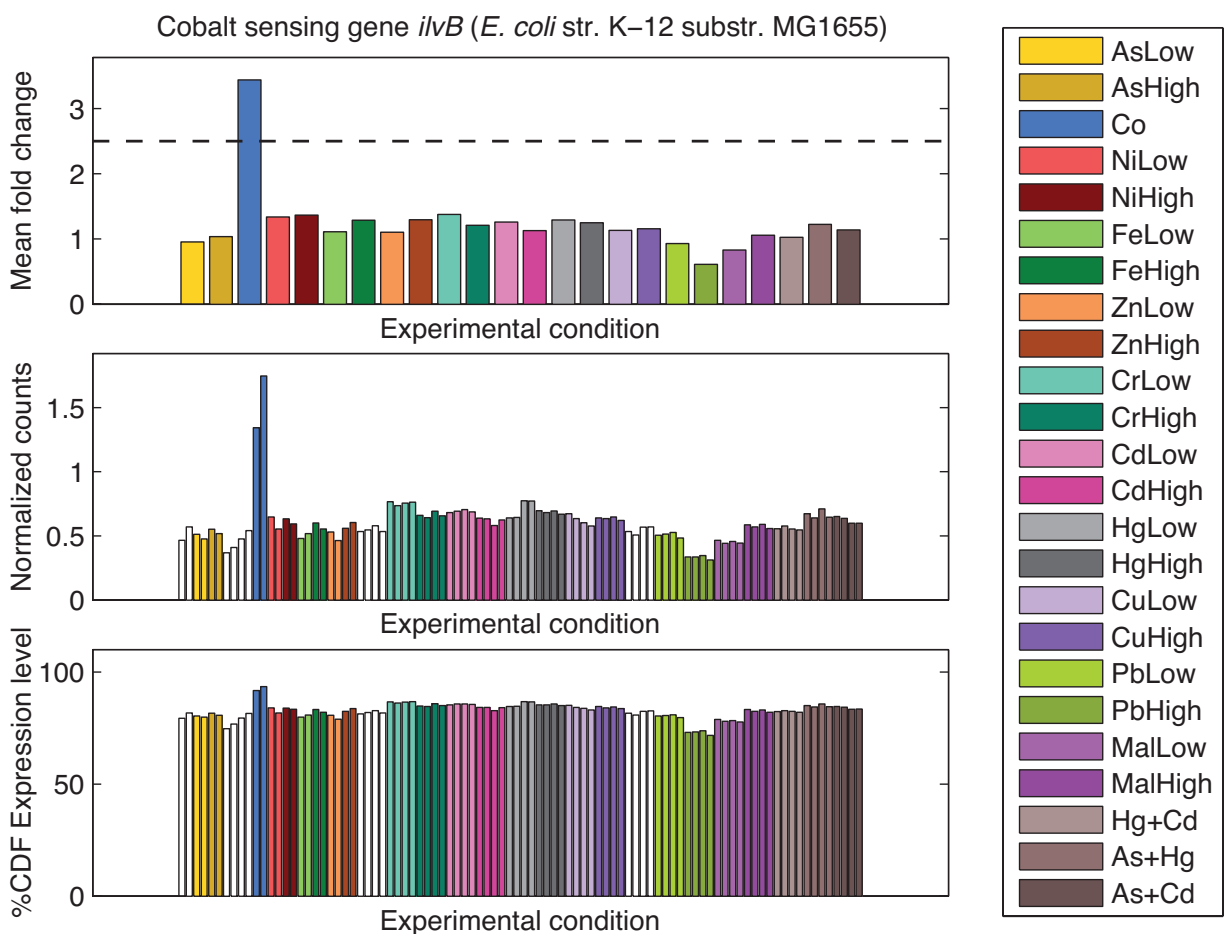
*Figure 35: Response of the ilvB promoter in* **E. coli** *MG1655 for specifically sensing cobalt. Note that the response of the ilvB promoter is specific to cobalt alone.*
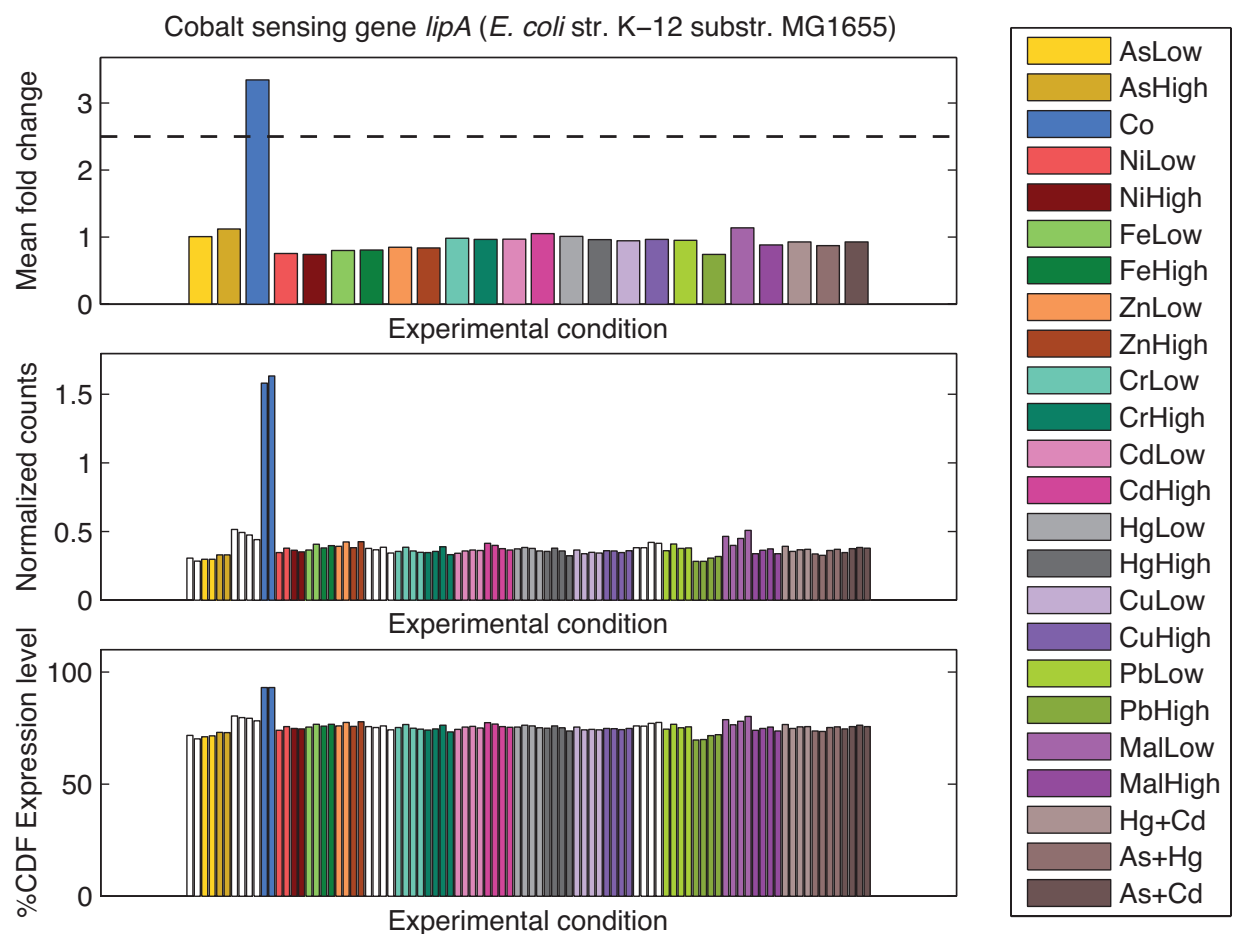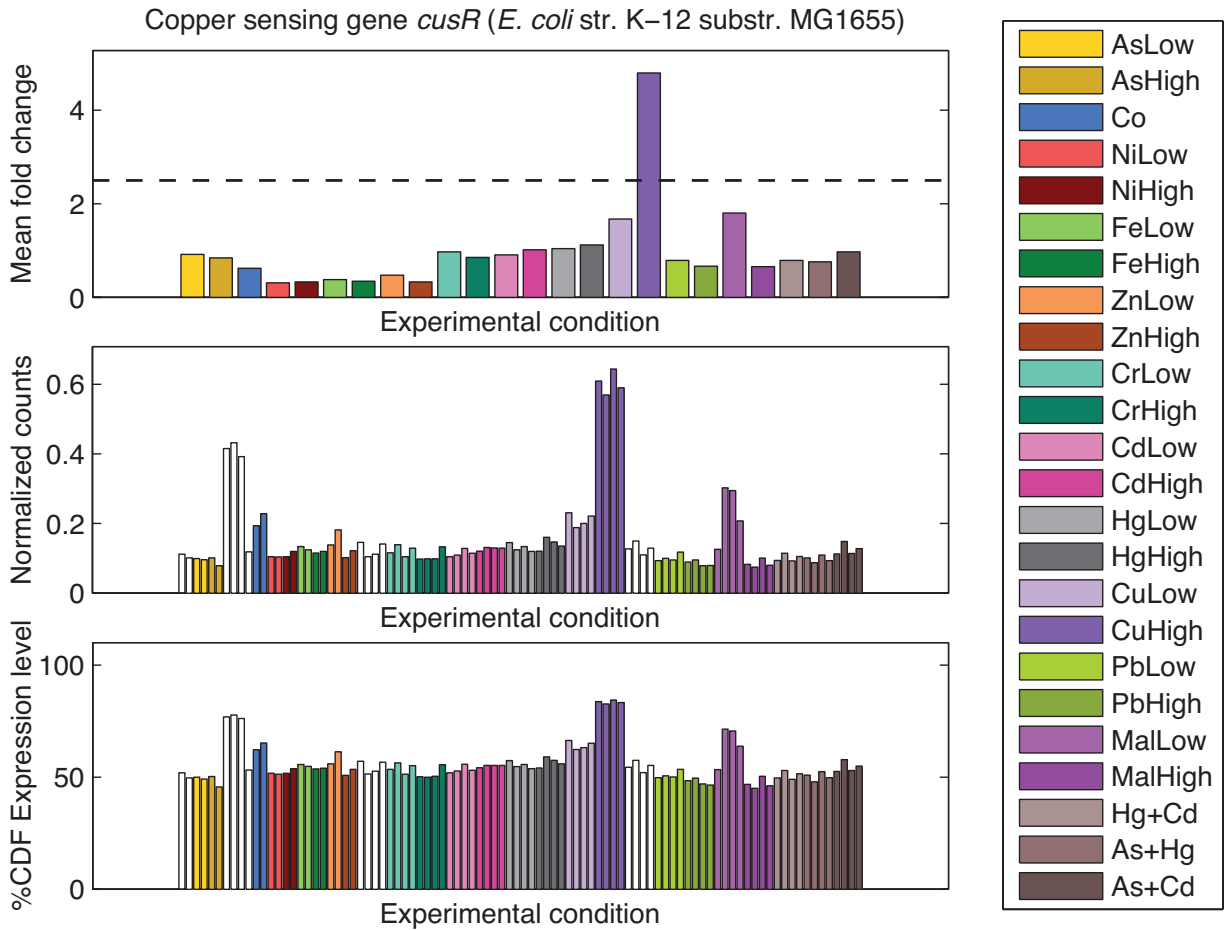
*Figure 36: Response of the lipA promoter in* **E. coli** *MG1655 for specifically sensing cobalt. Note that the response of the lipA promoter is specific to cobalt alone.*
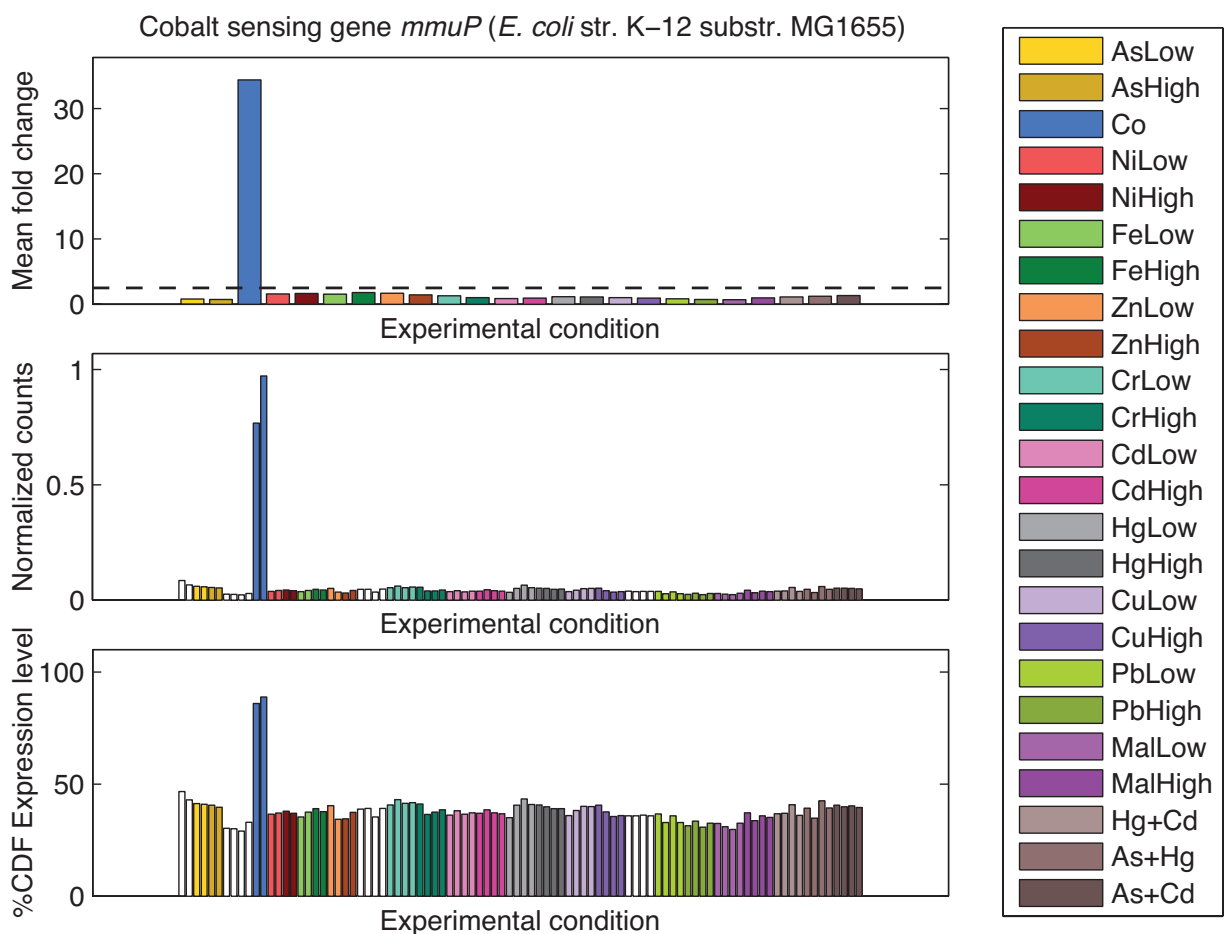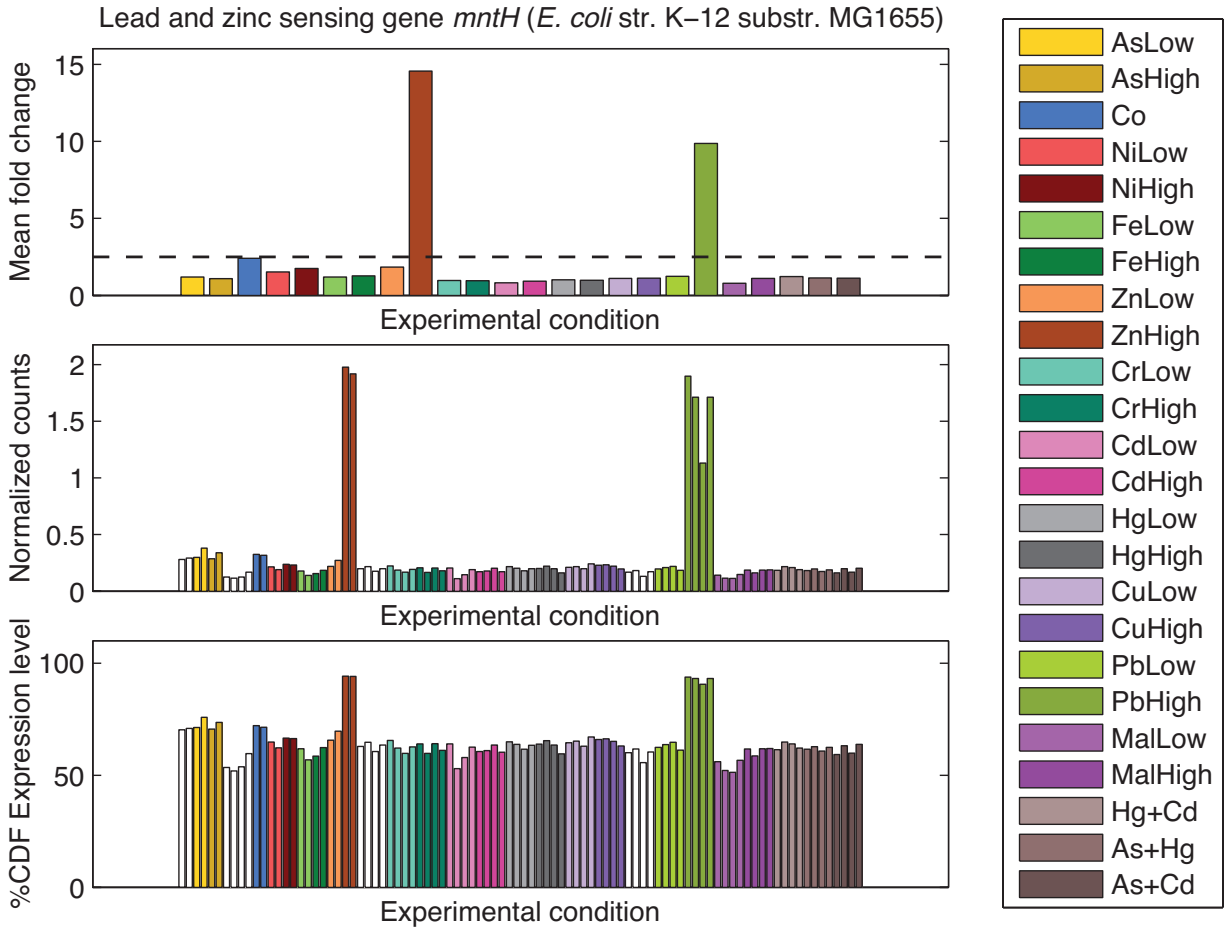
*Figure 37: Response of the mmuP promoter in* **E. coli** *MG1655 for specifically sensing cobalt. Note that the response of the mmuP promoter is specific to cobalt alone.*
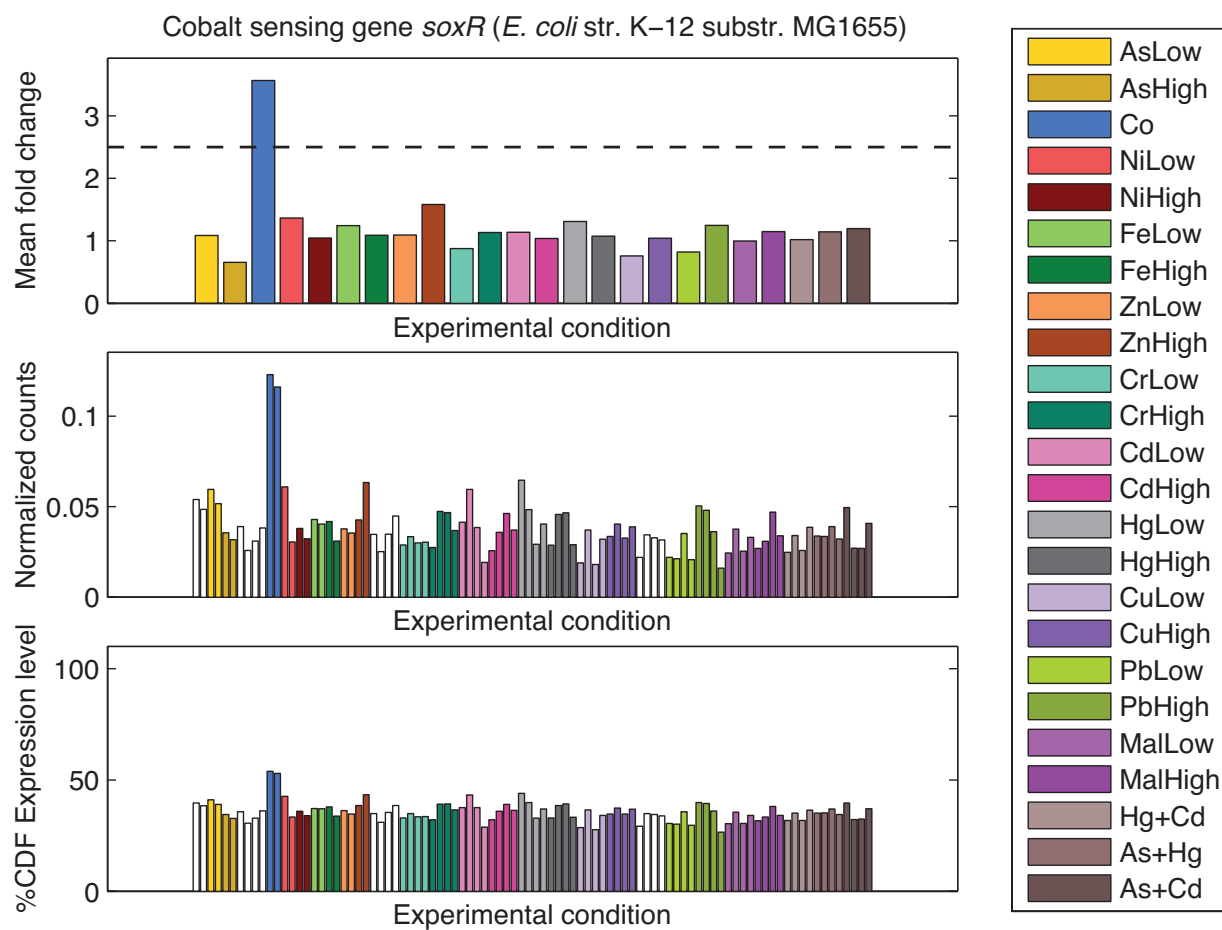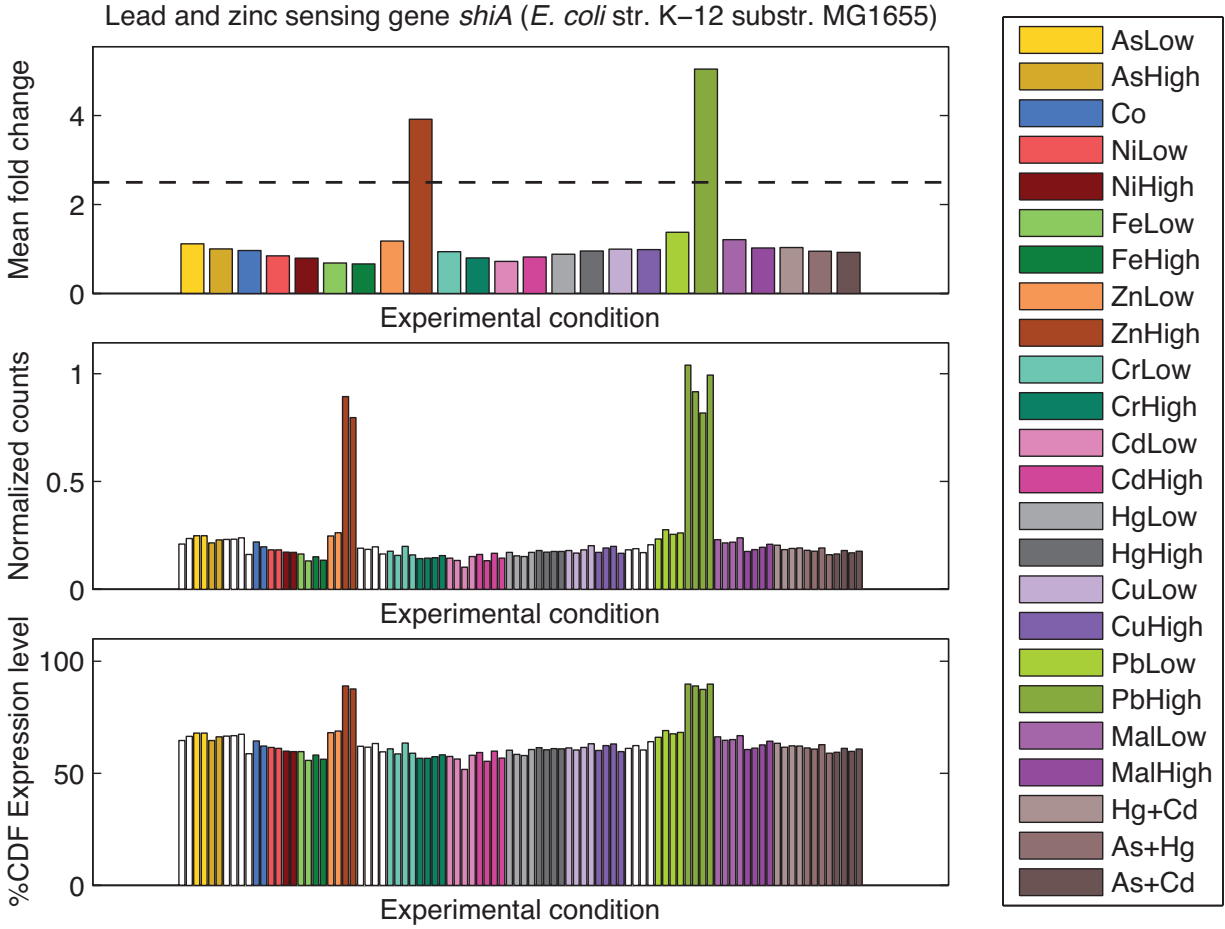
*Figure 38: Response of the soxR promoter in* **E. coli** *MG1655 for specifically sensing cobalt. Note that the response of the soxR promoter is specific to cobalt alone.*

*Figure 39: Response of the tehA promoter in* **E. coli** *MG1655 for specifically sensing cobalt. Note that the response of the tehA promoter is specific to cobalt alone.*

*Figure 40: Response of the ygbA promoter in* **E. coli** *MG1655 for specifically sensing cobalt. Note that the response of the ygbA promoter is specific to cobalt alone.*
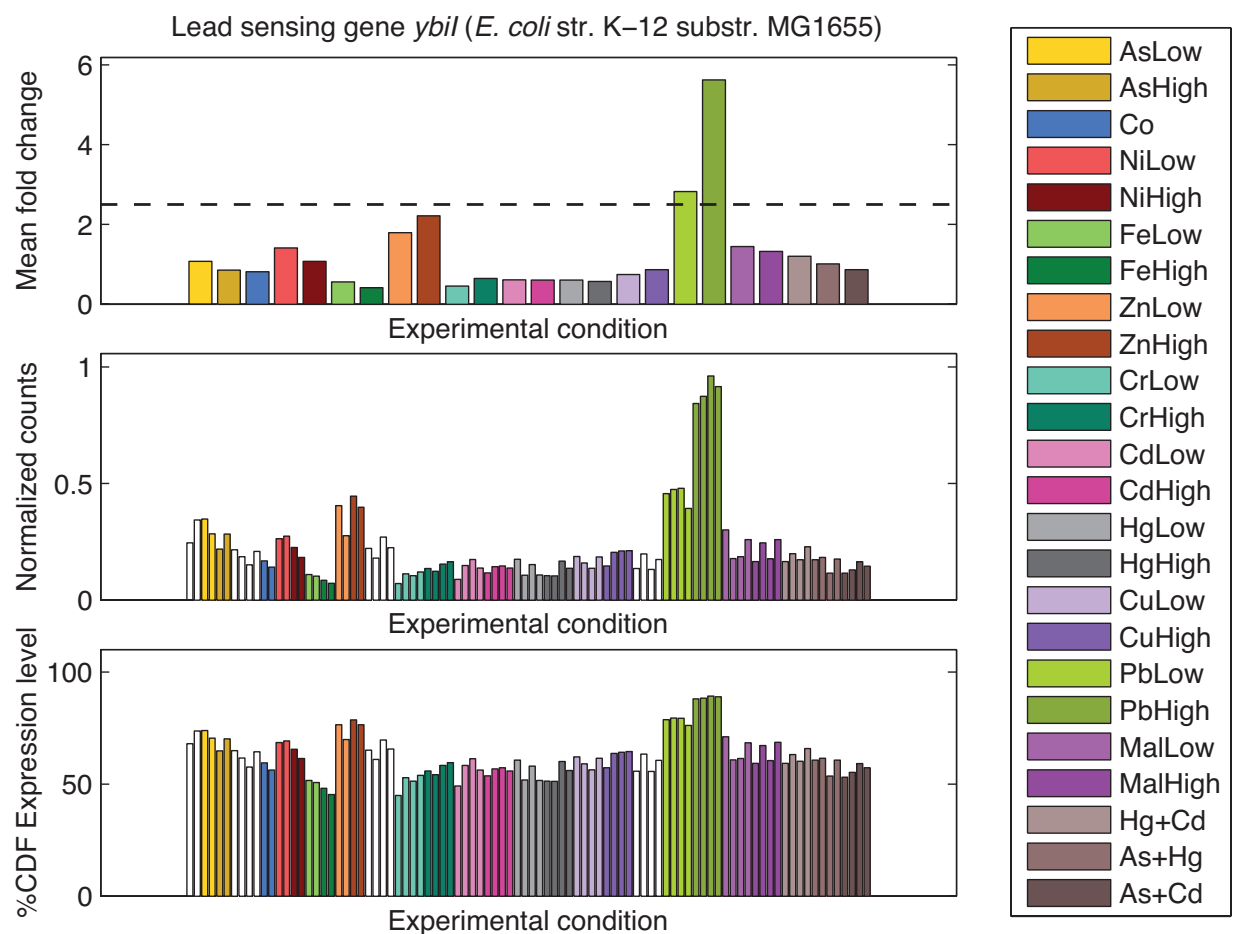
*Figure 41: Response of the yjbJ promoter in* **E. coli** *MG1655 for specifically sensing cobalt. Note that the response of the yjbJ promoter is specific to cobalt alone.*

*Figure 42: Response of the yqfA promoter in* **E. coli** *MG1655 for specifically sensing cobalt. Note that the response of the yqfA promoter is specific to cobalt alone.*

*Figure 43: Response of the cusC promoter in* **E. coli** *MG1655 for sensing copper and malathion. Note that the cusC promoter responds monotonically to increasing concentrations of copper alone and to low concentrations of malathion alone. A copper-specific sensor can be implemented by combining this response with the malathion-specific response of nemR using the boolean expression (cusC)∧(¬nemR). Alternatively, a malathion-specific sensor can be implemented by combining this response with the copper-specific response of cusR using the boolean expression (cusC)∧(¬cusR).*

*Figure 44: Response of the cusR promoter in* **E. coli** *MG1655 for specifically sensing copper. Note that the response of the cusR promoter is specific to copper alone but only shows sensitivity to high concentrations of copper.*
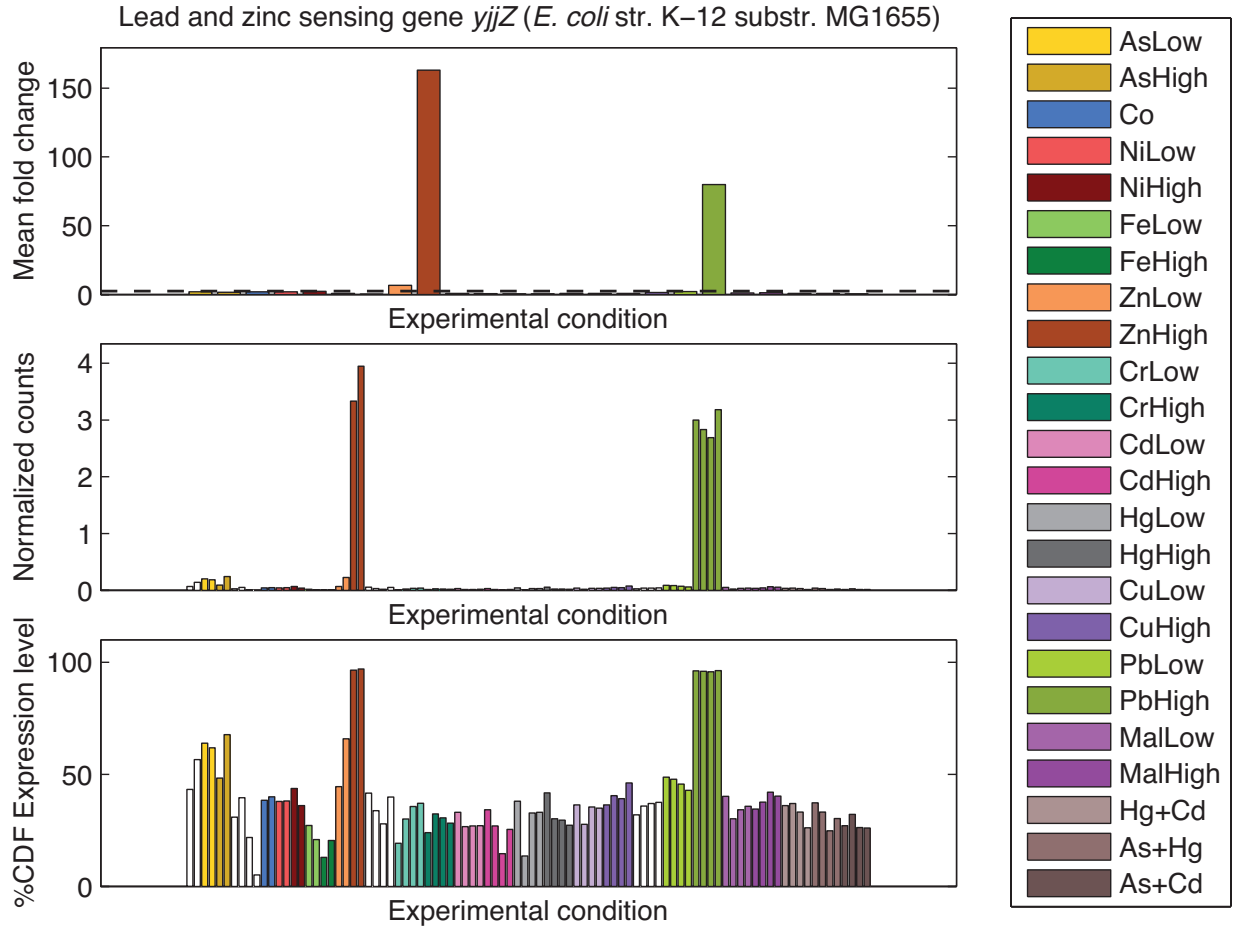
*Figure 45: Response of the mntH promoter in* **E. coli** *MG1655 for sensing lead and zinc. Note that the mntH promoter responds to high concentrations of lead alone and high concentrations of zinc alone. A lead-specific sensor can be implemented by combining this response with the zinc-specific response of zraP using the boolean expression (mntH)∧(¬zraP). Alternatively, a zinc-specific sensor can be implemented by combining this response with the lead-specific response of ybiI using the boolean expression (mntH)∧(¬ybiI).*

*Figure 46: Response of the shiA promoter in* **E. coli** *MG1655 for sensing lead and zinc. Note that the shiA promoter responds to high concentrations of lead alone and high concentrations of zinc alone. A lead-specific sensor can be implemented by combining this response with the zinc-specific response of zraP using the boolean expression (shiA)∧(¬zraP). Alternatively, a zinc-specific sensor can be implemented by combining this response with the lead-specific response of ybiI using the boolean expression (shiA)∧(¬ybiI).*
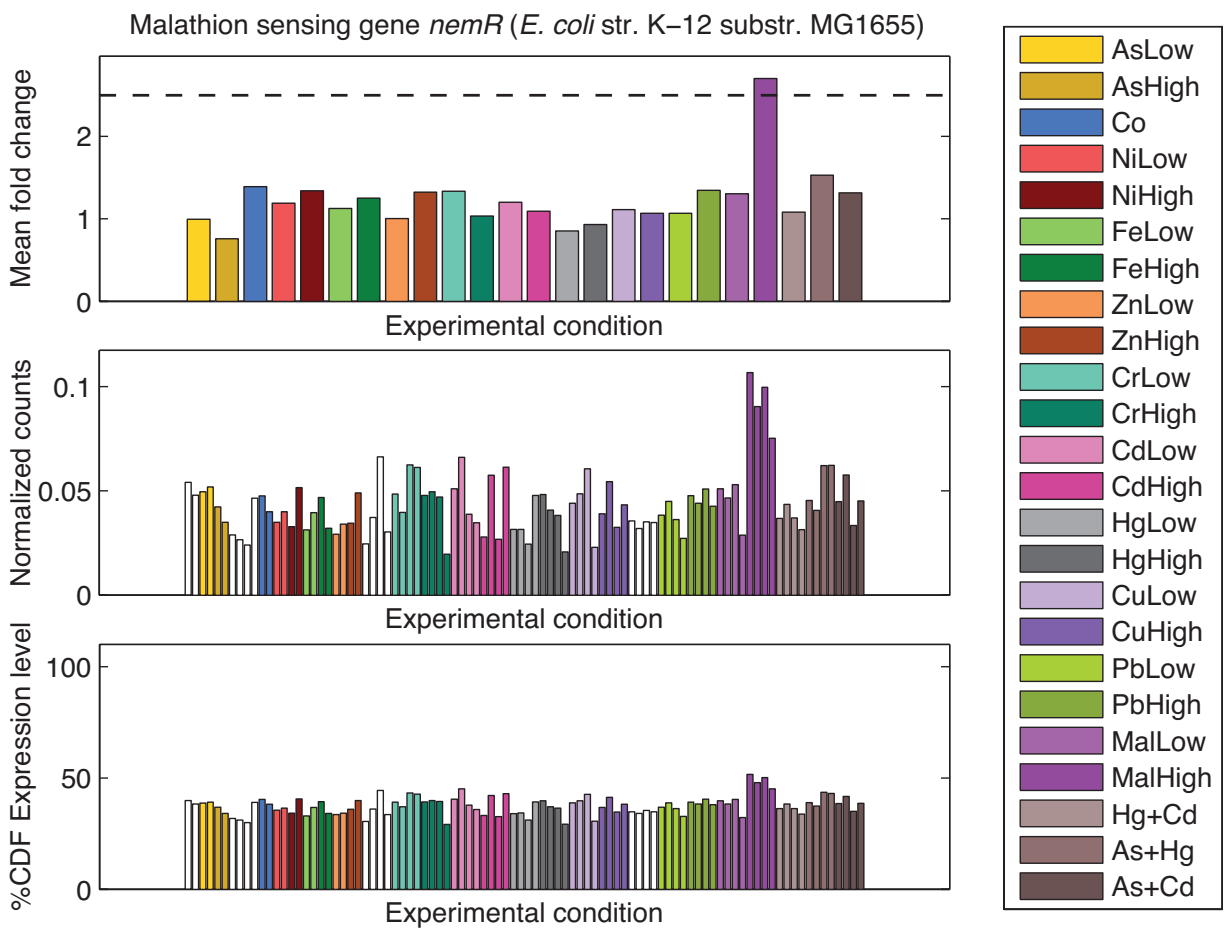
*Figure 47: Response of the ybiI promoter in* **E. coli** *MG1655 for specifically sensing lead. Note that the ybiI promoter responds specifically and monotonically to increasing concentrations of lead alone.*
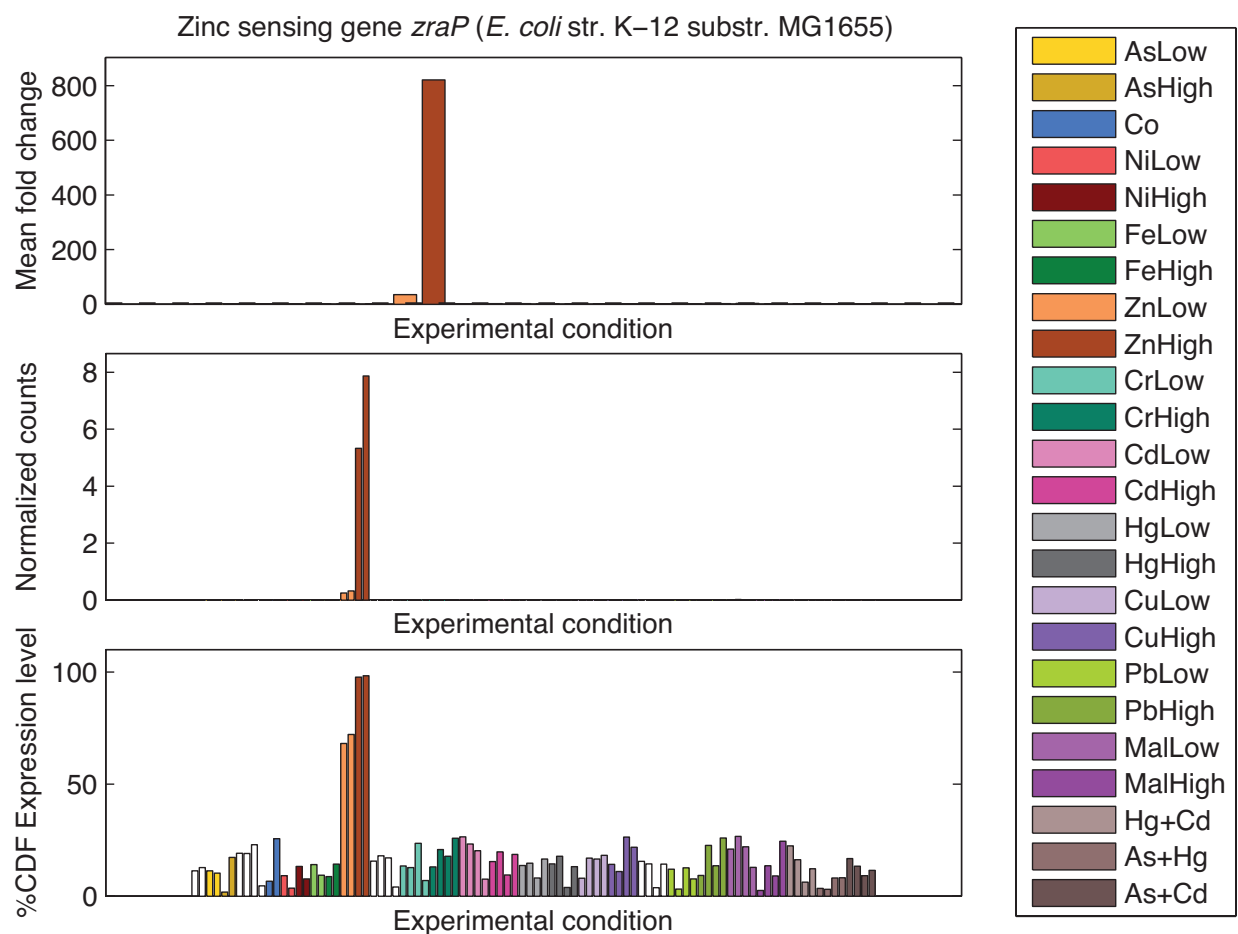
*Figure 48: Response of the yjjZ promoter in* **E. coli** *MG1655 for sensing lead and zinc. Note that the yjjZ promoter responds monotonically to increasing concentrations of zinc alone and to high concentrations of lead alone. A lead-specific sensor can be implemented by combining this response with the zinc-specific response of zraP using the boolean expression (yjjZ)∧(¬zraP). Alternatively, a zinc-specific sensor can be implemented by combining this response with the lead-specific response of ybiI using the boolean expression (yjjZ)∧(¬ybiI).*

*Figure 49: Response of the nemR promoter in* **E. coli** *MG1655 for specifically sensing malathion. Note that the nemR promoter responds to high concentrations of malathion alone.*

*Figure 50: Response of the zraP promoter in* **E. coli** *MG1655 for specifically sensing zinc. Note that the zraP promoter responds specifically and monotonically to increasing concentrations of zinc alone.*